SPIS TREŚCI.

I GEOMETRIA ANALITYCZNA W PRZESTRZENI 1
1. Algebra wektorów ... 1
2. Płaszczyzna ... 18
3. Prosta w przestrzeni ... 31
4. Prosta i płaszczyzna ... 40
5. Powierzchnie ... 51

II RACHUNEK RÓŻNICZKOWY FUNKCJI WIELU ZMIENNYCH 65
1. Przestrzeń metryczna i pewne własności jej podzbiorów ... 65
2. Zbieżność ciągu w przestrzeni metrycznej ... 75
3. Funkcje wielu zmiennych ... 79
4. Granica i ciągłość funkcji wielu zmiennych ... 89
5. Pochodne cząstkowe. Różniczka funkcji ... 98
6. Ekstrema funkcji dwu zmiennych ... 114
7. Funkcja uwikłana ... 128

III RACHUNEK CAŁKOWY FUNKCJI WIELU ZMIENNYCH 136
1. Określenie całki podwójnej i jej interpretacja geometryczna ... 136
2. Własności i obliczanie całki podwójnej ... 143
3. Zamiana zmiennych w całce podwójnej ... 153
4. Zastosowania geometryczne całki podwójnej ... 164
5. Przykłady zastosowania całki podwójnej w fizyce ... 172
6. Całka krzywoliniowa nieskierowana ... 175

7. Całka krzywoliniowa skierowana 184
8. Całka potrójna 206

IV RÓWNANIA RÓŻNICZKOWE ZWYCZAJNE 214
1. Podstawowe wiadomości 214
2. Równania o zmiennych rozdzielonych. Równanie jednorodne 228
3. Równanie liniowe pierwszego rzędu. Równanie Bernoulliego 245
4. Równanie zupełne. Czynnik całkujący 262
5. Równania różniczkowe drugiego rzędu sprowadzalne do równań pierwszego rzędu 269
6. Równanie liniowe drugiego rzędu 277
7. Pewne wiadomości o układach równań różniczkowych 302

V ELEMENTY RACHUNKU PRAWDOPODOBIEŃSTWA 312
1. Zdarzenia losowe 312
2. Prawdopodobieństwo 317
3. Zmienne losowe 330
4. Charakterystyki liczbowe zmiennych losowych 349
5. Niektóre skokowe rozkłady pradopodobieństwa 363
6. Niektóre ciągłe rozkłady prawdopodobieństwa 369
7. Wektory losowe 377

VI ELEMENTY STATYSTYKI MATEMATYCZNEJ 399
1. Trzy rozkłady ważne w statystyce 399
2. Przykłady twierdzeń granicznych 405
3. Próba. Statystyki. Próbka i jej opis 411
4. Estymatory punktowe 416
5. Estymacja przedziałowa 421
6. Wstępne pojęcia weryfikacji hipotez 428
7. Testy dla wartości oczekiwanej 433
8. Test zgodności chi-kwadrat 439

TABLICE STATYSTYCZNE 445

SKOROWIDZ 449

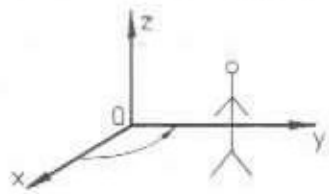
I. GEOMETRIA ANALITYCZNA W PRZESTRZENI

1. ALGEBRA WEKTORÓW.

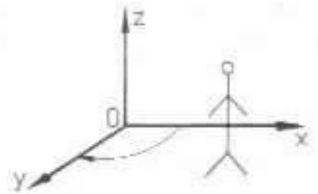
PODSTAWOWE WIADOMOŚCI O WEKTORACH. Przypomnimy znane ze szkoły średniej podstawowe pojęcia i wzory dotyczące wektorów w R^3 z prostokątnym układem współrzędnych 0xyz.

Prostokątnym układem współrzędnych 0xyz nazywamy uporządkowaną trójkę osi 0x, 0y, 0z wzajemnie prostopadłych o wspólnym początku 0 i wspólnej jednostce.

Wyróżnia się dwa rodzaje układów współrzędnych 0xyz: **prawoskrętny** (rys 1.1.) i **lewoskrętny** (rys 1.2.).



Rys 1.1.

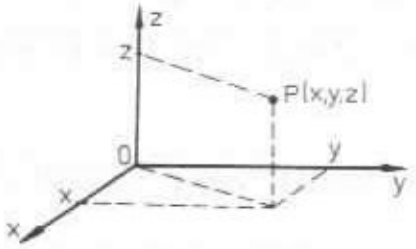


Rys 1.2.

Prostokątny układ współrzędnych 0xyz jest prawoskrętny, gdy dla obserwatora umieszczonego tak, jak na rysunkach 1.1 i 1.2 obrót osi 0x do pokrycia się z osią 0y o kąt $\pi/2$ jest obrotem w kierunku przeciwnym do ruchu wskazówek zegara. W przeciwnym przypadku układ współrzędnych 0xyz jest lewoskrętny.

Układ współrzędnych prawo- lub lewoskrętny nazywamy **układem zorientowanym**.

Każdemu punktowi $P \in R^3$ przyporządkowujemy uporządkowaną trójkę liczb x, y, z zwanych **współrzędnymi** tego **punktu** P i odwrotnie - każdej trójce liczb x, y, z przyporządkowujemy dokładnie jeden punkt P (rys 1.3) i zapisujemy wówczas $P(x,y,z)$ lub $P=(x,y,z)$; przy czym x jest współrzędną rzutu prostokątnego punktu P na oś 0x, analogicznie y i z są współrzędnymi rzutu prostokątnego punktu P odpowiednio na oś 0y i 0z.



Rys 1.3.

Wektorem nazywamy uporządkowaną parę punktów P, Q; pierwszy z tych punktów, punkt P, nazywamy **początkiem**, drugi, punkt Q- **końcem** wektora. Wektor taki oznaczamy symbolem $\overrightarrow{PQ}$.

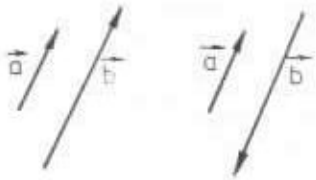
Wektorem zerowym nazywamy wektor, którego początek pokrywa się z końcem i oznaczamy symbolem $\vec{0}$.

Wektory oznaczamy również jedną literą np. $\vec{a}$, $\vec{b}$, $\vec{w}$, $\vec{r}$.

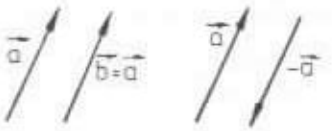
Długością wektora $\overrightarrow{PQ}$, oznaczaną symbolem $|\overrightarrow{PQ}|$ lub PQ, nazywamy długość odcinka $\overline{PQ}$; długość wektora $\vec{a}$ oznacza się symbolem $|\vec{a}|$ lub a.

Kierunkiem niezerowego wektora $\overrightarrow{PQ}$ nazywamy kierunek prostej PQ, **zwrotem** wektora $\overrightarrow{PQ}$ nazywamy ten z dwu zwrotów prostej PQ, w którym punkt P poprzedza punkt Q. Dla wektora zerowego nie określamy kierunku i zwrotu.

Wersorem osi nazywamy wektor o kierunku i zwrocie tej osi i o długości równej 1; wersory osi 0x, 0y, 0z oznaczamy odpowiednio $\vec{i}$, $\vec{j}$, $\vec{k}$.



Rys 1.4.



Rys 1.5

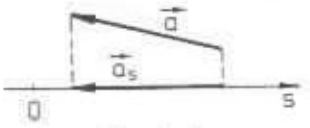
Dwa niezerowe wektory $\vec{a}$ i $\vec{b}$ nazywamy **równoległymi**, co zapisujemy $\vec{a} \parallel \vec{b}$, gdy mają ten sam kierunek. Dwa wektory równoległe mogą mieć **zwroty zgodne** lub **zwroty przeciwne** (rys 1.4).

Dwa niezerowe wektory $\vec{a}$ i $\vec{b}$ nazywamy **równymi**, co zapisujemy $\vec{a}=\vec{b}$, gdy mają ten sam kierunek, zwrot i tę samą długość (rys 1.5). Dodatkowo przyjmujemy, że dwa dowolne wektory zerowe są równe.

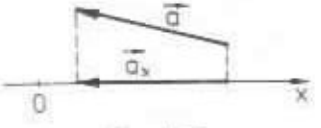
Dwa wektory nazywamy **przeciwnymi**, gdy mają ten sam kierunek, tę samą długość i zwroty przeciwne; wektor przeciwny do $\vec{a}$ oznaczamy symbolem $-\vec{a}$ (rys 1.5).

Mówimy, że wektor $\vec{a}$ i oś s równoległa do niego mają zwroty zgodne (przeciwne), gdy wektor $\vec{a}$ i wersor tej osi mają zwroty zgodne (przeciwne).

Rzut prostokątny wektora $\vec{a}$ na oś s jest wektorem i oznacza się go symbolem $\vec{a}_s$ (rys 1.6).



Rys 1.6.



Rys 1.7.

Współrzędną wektora $\vec{a}$ względem osi s, oznaczaną symbolem a_s, nazywamy liczbę określoną wzorem

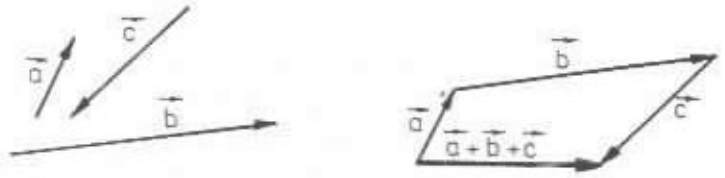
$$a_s=\begin{cases} |\vec{a}_s|, & \text{gdy } \vec{a}_s, s \text{ mają zwroty zgodne,} \\ -|\vec{a}_s|, & \text{gdy } \vec{a}_s, s \text{ mają zwroty przeciwne,} \\ 0, & \text{gdy } \vec{a}_s=\vec{0}. \end{cases}$$

W szczególności, gdy s jest osią 0x, wektor $\vec{a}_x$ jest rzutem prostokątnym wektora $\vec{a}$ na oś 0x, a_x - współrzędną wektora $\vec{a}$ względem

osi 0x (rys 1.7). Wektor $\bar{a}$ ma w układzie współrzędnych 0xyz trzy współrzędne a_x, a_y, a_z, co zapisujemy również $\bar{a} = [a_x, a_y, a_z]$.

Sumę kilku wektorów ilustrujemy na rysunku (rys 1.8).

Różnicą dwu wektorów $\bar{a}$ i $\bar{b}$ nazywamy sumę: wektora $\bar{a}$ oraz wektora $-\bar{b}$.



Rys 1.8.

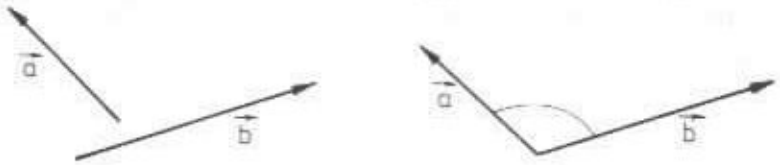
Iloczynem niezerowego **wektora** $\bar{a}$ **przez liczbę** $\alpha \neq 0$ nazywamy wektor $\bar{w}$, $\bar{w} = \alpha\bar{a}$, który:

1) jest równoległy do wektora $\bar{a}$, $\bar{w} \parallel \bar{a}$,

2) ma zwrot zgodny z wektorem $\bar{a}$, gdy $\alpha > 0$ lub zwrot przeciwny z wektorem $\bar{a}$, gdy $\alpha < 0$,

3) ma długość $|\bar{w}| = |\alpha| \cdot |\bar{a}|$.

Dodatkowo przyjmujemy: $\alpha \cdot \bar{a} = \bar{0}$, gdy $\alpha = 0$ lub $\bar{a} = \bar{0}$.



Rys 1.9.

Kąt (nieskierowany) wektorów niezerowych $\bar{a}$ i $\bar{b}$ (kąt między wektorami) oznaczany symbolem $\measuredangle(\bar{a}, \bar{b})$ jest figurą wypukłą, przy czym $\measuredangle(\bar{a}, \bar{b}) = \measuredangle(\bar{b}, \bar{a})$. Miara tego kąta jest liczbą z przedziału $<0, \pi>$.

Kątem wektora z osią nazywamy kąt tego wektora z wersorem tej osi np. $\measuredangle(\bar{a}, 0x) = \measuredangle(\bar{a}, \bar{i})$.

Umieścimy teraz wektory w układzie współrzędnych 0xyz. Jeśli punkty $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$ są odpowiednio początkiem i końcem

wektora $\bar{a} = \overrightarrow{P_1P_2}$, to współrzędne a_x, a_y, a_z i długość $|\bar{a}|$ tego wektora dane są wzorami

$$a_x = x_2 - x_1, \quad a_y = y_2 - y_1, \quad a_z = z_2 - z_1,$$

$$|\bar{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}.$$

Łatwo obliczamy współrzędne wersorów osi układu współrzędnych:

$$\bar{i} = [1,0,0], \quad \bar{j} = [0,1,0], \quad \bar{k} = [0,0,1].$$

Z przyjętych wcześniej definicji wynika, że:

Jeżeli $\bar{a} = [a_x, a_y, a_z]$, $\bar{b} = [b_x, b_y, b_z]$, to

$$\bar{a} = \bar{b} \Leftrightarrow (a_x = b_x \wedge a_y = b_y \wedge a_z = b_z),$$

$$\bar{a} \pm \bar{b} = [a_x \pm b_x, a_y \pm b_y, a_z \pm b_z],$$

$$\alpha\bar{a} = [\alpha a_x, \alpha a_y, \alpha a_z],$$

$$\bar{a} = a_x\bar{i} + a_y\bar{j} + a_z\bar{k}.$$

ILOCZYN SKALARNY. **Iloczynem skalarnym** dwu niezerowych wektorów $\bar{a}$ i $\bar{b}$ nazywamy liczbę $\bar{a} \circ \bar{b}$ określoną wzorem

$$\bar{a} \circ \bar{b} \overset{\text{def}}{=} |\bar{a}| \cdot |\bar{b}| \cos\varphi, \quad \text{gdzie } \varphi = \measuredangle(\bar{a}, \bar{b}).$$

Iloczynem skalarnym dwu wektorów, z których przynajmniej jeden jest wektorem zerowym nazywamy liczbę równą zero.

Z definicji tej wynikają następujące własności iloczynu skalarnego:

(1) $\bar{a} \circ \bar{b} = \bar{b} \circ \bar{a}$ przemienność,

(2) $\bar{a} \circ (\bar{b} + \bar{c}) = \bar{a} \circ \bar{b} + \bar{a} \circ \bar{c}$ rozdzielność względem dodawania,

(3) $\alpha(\bar{a} \circ \bar{b}) = (\alpha\bar{a}) \circ \bar{b} = \bar{a} \circ (\alpha\bar{b})$ dla $\alpha \in R$,

(4) $\bar{i} \circ \bar{i} = 1$, $\bar{j} \circ \bar{j} = 1$, $\bar{k} \circ \bar{k} = 1$,
$\bar{i} \circ \bar{j} = 0$, $\bar{i} \circ \bar{k} = 0$, $\bar{j} \circ \bar{k} = 0$.

TWIERDZENIE 1.1. (podstawowe o iloczynie skalarnym). Jeżeli

$$\bar{a} = [a_x, a_y, a_z], \quad \bar{b} = [b_x, b_y, b_z],$$

to iloczyn skalarny $\bar{a} \circ \bar{b}$ wyraża się wzorem

(1.1) $$\bar{a} \circ \bar{b} = a_x b_x + a_y b_y + a_z b_z.$$

D o w ó d . Wektory $\vec{a}$, $\vec{b}$ zapisujemy w postaci

$$\vec{a}=a_x\vec{i}+a_y\vec{j}+a_z\vec{k},\quad \vec{b}=b_x\vec{i}+b_y\vec{j}+b_z\vec{k}.$$

Następnie korzystając z własności iloczynu skalarnego otrzymujemy

$$\vec{a}\circ\vec{b}=(a_x\vec{i}+a_y\vec{j}+a_z\vec{k})\circ(b_x\vec{i}+b_y\vec{j}+b_z\vec{k})=\{\text{własn. (2) i (3)}\}=$$
$$=a_xb_x(\vec{i}\circ\vec{i})+a_xb_y(\vec{i}\circ\vec{j})+a_xb_z(\vec{i}\circ\vec{k})+$$
$$+a_yb_x(\vec{j}\circ\vec{i})+a_yb_y(\vec{j}\circ\vec{j})+a_yb_z(\vec{j}\circ\vec{k})+$$
$$+a_zb_x(\vec{k}\circ\vec{i})+a_zb_y(\vec{k}\circ\vec{j})+a_zb_z(\vec{k}\circ\vec{k})=\{\text{własn. (4)}\}=$$
$$=a_xb_x+a_yb_y+a_zb_z.$$ □

TWIERDZENIE 1.2. Niech $\vec{a}\neq\vec{0}$, $\vec{b}\neq\vec{0}$ oraz $\vec{a}=[a_x,a_y,a_z]$, $\vec{b}=[b_x,b_y,b_z]$. Wówczas

(1) $\vec{a}\perp\vec{b}\Leftrightarrow a_xb_x+a_yb_y+a_zb_z=0$,

(2) $$\cos\measuredangle(\vec{a},\vec{b})=\frac{a_xb_x+a_yb_y+a_zb_z}{\sqrt{a_x^2+a_y^2+a_z^2}\cdot\sqrt{b_x^2+b_y^2+b_z^2}}.$$

D o w ó d . Przy założeniu, że $\vec{a}\neq\vec{0}$, $\vec{b}\neq\vec{0}$ mamy

$$\vec{a}\perp\vec{b}\Leftrightarrow\vec{a}\circ\vec{b}=0.$$

Stąd i z twierdzenia 1.1 otrzymujemy warunek

$$\vec{a}\perp\vec{b}\Leftrightarrow a_xb_x+a_yb_y+a_zb_z=0.$$

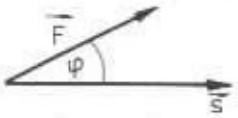
Z definicji iloczynu skalarnego wektorów niezerowych mamy:

$$\cos\measuredangle(\vec{a},\vec{b})=\frac{\vec{a}\circ\vec{b}}{|\vec{a}|\cdot|\vec{b}|}.$$

Uwzględniając w powyższym wzór (1.1) oraz wzory na długość wektorów otrzymamy natychmiast tezę (2). □

Przypomnijmy jeszcze interpretację fizyczną iloczynu skalarnego:

praca W, jaką wykona stała siła $\vec{F}$ przesuwając punkt materialny na prostoliniowej drodze $\vec{s}$, jest równa iloczynowi skalarnemu $\vec{F}\circ\vec{s}$:



$$W=|\vec{F}|\cdot|\vec{s}|\cos\varphi=\vec{F}\circ\vec{s}.$$

Na przykład siła $\vec{F}=4\vec{i}+3\vec{j}+\vec{k}$ na drodze $\vec{s}=2\vec{i}-\vec{j}+2\vec{k}$ wykonuje pracę

$$W=\vec{F}\circ\vec{s}=[4,3,1]\circ[2,-1,2]=4\cdot2+3\cdot(-1)+1\cdot2=7.$$

Prostokątny rzut wektora $\vec{a}$ na oś wyznaczoną przez wektor $\vec{b}$ jest wektorem i oznacza się go symbolem $\vec{a}_{\vec{b}}$, zaś współrzędna wektora $\vec{a}$ względem osi wyznaczonej przez wektor $\vec{b}$ jest liczbą, oznaczoną symbolem $a_{\vec{b}}$, określoną wzorem

$$a_{\vec{b}}=\begin{cases}|\vec{a}_{\vec{b}}|, & \text{gdy } \vec{a}_{\vec{b}},\vec{b} \text{ mają zwroty zgodne,}\\ -|\vec{a}_{\vec{b}}|, & \text{gdy } \vec{a}_{\vec{b}},\vec{b} \text{ mają zwroty przeciwne,}\\ 0, & \text{gdy } \vec{a}_{\vec{b}}=\vec{0}.\end{cases}$$

Rys 1.10.

Można wykazać, że współrzędna $a_{\vec{b}}$ (rys 1.10) wyraża się wzorem

$$a_{\vec{b}}=|\vec{a}|\cos\measuredangle(\vec{a},\vec{b})=|\vec{a}|\cos\varphi.$$

W konsekwencji

$$\vec{a}\circ\vec{b}=a_{\vec{b}}\cdot|\vec{b}|=|\vec{a}|\cdot b_{\vec{a}}.$$

P R Z Y K Ł A D 1.1. Obliczymy współrzędną $a_{\vec{b}}$, gdy $\vec{a}=[-5,1,-2]$, $\vec{b}=[0,8,6]$.

Współrzędna ta jest równa

$$a_{\vec{b}}=|\vec{a}|\cos\measuredangle(\vec{a},\vec{b})=|\vec{a}|\cdot\frac{\vec{a}\circ\vec{b}}{|\vec{a}|\cdot|\vec{b}|}=\frac{\vec{a}\circ\vec{b}}{|\vec{b}|}=-\frac{2}{5},$$

gdyż $\vec{a}\circ\vec{b}=[-5,1,-2]\circ[0,8,6]=-5\cdot0+1\cdot8-2\cdot6=-4$,

$|\vec{b}|=\sqrt{0^2+8^2+6^2}=10$.

Zauważmy jeszcze: z tego, że współrzędna ta jest ujemna wynika, że kąt między wektorami $\bar{a}$ i $\bar{b}$ jest rozwarty, ponadto długość prostokątnego rzutu wektora $\bar{a}$ na oś wektora $\bar{b}$ jest równa 2/5. ■

PRZYKŁAD 1.2. Dobierzemy parametr m tak, by wektory [m,3,4], [m,m,−1] były prostopadłe.

Wektory te są prostopadłe wtedy i tylko wtedy, gdy ich iloczyn skalarny jest równy zeru:

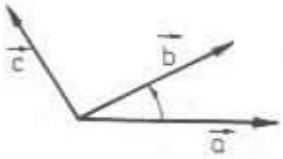
$$[m,3,4]\circ[m,m,-1]=0 \Leftrightarrow m^2+3m-4=0 \Leftrightarrow (m=1 \vee m=-4).$$

Zatem wektory te są prostopadłe dla $m=1$ lub $m=-4$. ■

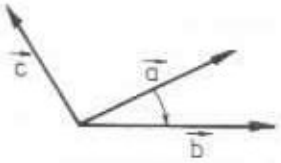
ILOCZYN WEKTOROWY. Iloczyn wektorowy jest związany ściśle z trójwymiarowym układem współrzędnych i jego orientacją oraz orientacją trójki wektorów.

Uporządkowana trójka wektorów $\bar{a},\bar{b},\bar{c}$ o wspólnym początku, niezerowych i nierównoległych do jednej płaszczyzny, z których żadne dwa wektory nie są równoległe, może mieć orientację zgodną lub przeciwną do orientacji układu współrzędnych 0xyz.

O uporządkowanej trójce wektorów $\bar{a},\bar{b},\bar{c}$ z rysunku 1.11 mówimy, że ma **orientację zgodną** z prawoskrętnym układem współrzędnych 0xyz, zaś uporządkowana trójka wektorów $\bar{a},\bar{b},\bar{c}$ z rysunku 1.12 ma **orientację przeciwną** do prawoskrętnego układu współrzędnych 0xyz, (a zgodną z lewoskrętnym układem współrzędnych 0xyz).



Rys 1.11.



Rys 1.12.

Oczywiście uporządkowana trójka wersorów $\bar{i},\bar{j},\bar{k}$ ma orientację zgodną ze "swoim" zorientowanym układem współrzędnych 0xyz; taką samą orientację mają również trójki wersorów

$$\bar{j},\bar{k},\bar{i} \qquad \bar{k},\bar{i},\bar{j}.$$

Natomiast orientację przeciwną ze "swoim" zorientowanym układem współrzędnych 0xyz mają trójki wersorów

$$\bar{i},\bar{k},\bar{j}, \qquad \bar{j},\bar{i},\bar{k}, \qquad \bar{k},\bar{j},\bar{i}.$$

Iloczynem wektorowym uporządkowanej pary niezerowych i nierównoległych wektorów $\bar{a}$ i $\bar{b}$ w układzie współrzędnych 0xyz nazywamy wektor $\bar{w}=\bar{a}\times\bar{b}$ spełniający warunki:

1) wektor $\bar{w}$ jest prostopadły do każdego z wektorów $\bar{a}$ i $\bar{b}$,

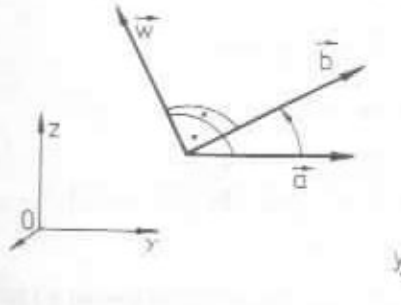
2) zwrot wektora $\bar{w}$ jest tak dobrany, by uporządkowana trójka wektorów $\bar{a},\bar{b},\bar{w}$ miała orientację zgodną z orientacją rozważanego układu współrzędnych 0xyz,

3) długość $|\bar{w}|$ wektora $\bar{w}$ jest określona wzorem

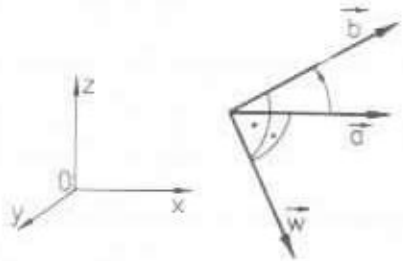
$$|\bar{w}|=|\bar{a}|\cdot|\bar{b}|\sin\varphi, \qquad \text{gdzie } \varphi=\sphericalangle(\bar{a},\bar{b}).$$

Iloczynem wektorowym dwu wektorów, z których przynajmniej jeden jest wektorem zerowym lub wektorów równoległych, nazywamy wektor zerowy.

Na rysunku 1.13 zaznaczono iloczyn wektorowy $\bar{w}=\bar{a}\times\bar{b}$ w prawoskrętnym układzie współrzędnych 0xyz zaś na rysunku 1.14 w lewoskrętnym układzie współrzędnych 0xyz.



Rys 1.13.



Rys 1.14.

Zauważmy, że długość iloczynu wektorowego ma następującą interpretację geometryczną:

długość iloczynu wektorowego wektorów $\bar{a}$ i $\bar{b}$ jest równa polu S równoległoboku "zbudowanego" na tych wektorach, czyli

$S = |\vec{a}| \cdot |\vec{b}| \sin\varphi = |\vec{a} \times \vec{b}|$.

Z definicji iloczynu wektorowego wynikają następujące jego własności:

(1) $\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$ antyprzemienność,

(2) $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$ rozdzielność względem dodawania,

(3) $\alpha(\vec{a} \times \vec{b}) = (\alpha\vec{a}) \times \vec{b} = \vec{a} \times (\alpha\vec{b})$, dla $\alpha \in R$,

(4) $\vec{i} \times \vec{j} = \vec{k}$, $\vec{j} \times \vec{k} = \vec{i}$, $\vec{k} \times \vec{i} = \vec{j}$,

$\vec{i} \times \vec{i} = \vec{0}$, $\vec{j} \times \vec{j} = \vec{0}$, $\vec{k} \times \vec{k} = \vec{0}$.

Dla przykładu wykażemy, że $\vec{i} \times \vec{j} = \vec{k}$. Iloczyn wektorowy $\vec{i} \times \vec{j}$ jest wektorem prostopadłym do wersora $\vec{i}$ oraz wersora $\vec{j}$, czyli ma kierunek osi 0z, ma zwrot osi 0z i długość równą $|\vec{i}| \cdot |\vec{j}| \sin\pi/2 = 1$. Zatem iloczyn $\vec{i} \times \vec{j}$ jest równy wersorowi osi 0z, czyli jest wektorem $\vec{k}$.

TWIERDZENIE 1.3 (podstawowe o iloczynie wektorowym). Jeżeli

$$\vec{a} = [a_x, a_y, a_z], \qquad \vec{b} = [b_x, b_y, b_z],$$

to iloczyn wektorowy $\vec{a} \times \vec{b}$ wyraża się wzorem

$$\vec{a} \times \vec{b} = \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} \vec{i} - \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} \vec{j} + \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} \vec{k}, \tag{1.2}$$

co można również zapisać w postaci

$$\vec{a} \times \vec{b} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}. \tag{1.3}$$

D o w ó d . Wektory $\vec{a}$, $\vec{b}$ zapisujemy w postaci

$$\vec{a} = a_x\vec{i} + a_y\vec{j} + a_z\vec{k}, \qquad \vec{b} = b_x\vec{i} + b_y\vec{j} + b_z\vec{k}.$$

Następnie korzystając z własności iloczynu wektorowego otrzymujemy:

$$\begin{aligned}
\vec{a} \times \vec{b} &= (a_x\vec{i} + a_y\vec{j} + a_z\vec{k}) \times (b_x\vec{i} + b_y\vec{j} + b_z\vec{k}) = \{\text{własn. (2) i (3)}\} = \\
&= a_xb_x(\vec{i} \times \vec{i}) + a_xb_y(\vec{i} \times \vec{j}) + a_xb_z(\vec{i} \times \vec{k}) + \\
&+ a_yb_x(\vec{j} \times \vec{i}) + a_yb_y(\vec{j} \times \vec{j}) + a_yb_z(\vec{j} \times \vec{k}) + \\
&+ a_zb_x(\vec{k} \times \vec{i}) + a_zb_y(\vec{k} \times \vec{j}) + a_zb_z(\vec{k} \times \vec{k}) = \{\text{własn. (4)}\} = \\
&= a_xb_y\vec{k} - a_xb_z\vec{j} - a_yb_x\vec{k} + a_yb_z\vec{i} + a_zb_x\vec{j} - a_zb_y\vec{i} = \\
&= (a_yb_z - a_zb_y)\vec{i} - (a_xb_z - a_zb_x)\vec{j} + (a_xb_y - a_yb_x)\vec{k} = \\
&= \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} \vec{i} - \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} \vec{j} + \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} \vec{k}.
\end{aligned}$$

□

TWIERDZENIE 1.4. Niech $\vec{a} \neq \vec{0}$, $\vec{b} \neq \vec{0}$ oraz $\vec{a} = [a_x, a_y, a_z]$, $\vec{b} = [b_x, b_y, b_z]$. Wówczas

(1) $\vec{a} \parallel \vec{b} \Leftrightarrow \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} = \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} = \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} = 0$,

(1') $\vec{a} \parallel \vec{b} \Leftrightarrow \dfrac{a_x}{b_x} = \dfrac{a_y}{b_y} = \dfrac{a_z}{b_z}$, przy założeniu, że $b_x, b_y, b_z \neq 0$,

(2) $\sin\measuredangle(\vec{a}, \vec{b}) = \dfrac{\left| \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} \right|}{\sqrt{a_x^2 + a_y^2 + a_z^2} \cdot \sqrt{b_x^2 + b_y^2 + b_z^2}}$.

D o w ó d . Ponieważ wektory $\vec{a}$ i $\vec{b}$ są niezerowe, więc

$$\vec{a} \parallel \vec{b} \Leftrightarrow \vec{a} \times \vec{b} = \vec{0}.$$

Stąd i z twierdzenia 1.3 otrzymujemy warunek

$$\vec{a} \parallel \vec{b} \Leftrightarrow \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} \vec{i} - \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} \vec{j} + \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} \vec{k} = \vec{0} \Leftrightarrow$$

$$\Leftrightarrow \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} = 0 \wedge \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} = 0 \wedge \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} = 0.$$

Jeśli $b_x \neq 0$, $b_y \neq 0$, $b_z \neq 0$, to warunek (1) można zapisać w łatwiejszej do zapamiętania postaci (1').

Z definicji długości iloczynu wektorowego wektorów niezerowych mamy

$$\sin \measuredangle(\vec{a},\vec{b}) = \frac{|\vec{a}\times\vec{b}|}{|\vec{a}|\cdot|\vec{b}|}.$$

Uwzględniając w powyższym wzór na długość wektorów otrzymujemy tezę (2). □

U w a g a . Przyjmijmy dodatkową umowę, że

$$\frac{a_x}{b_x} = \frac{a_y}{b_y} \Leftrightarrow a_x b_y = a_y b_x$$

tzn. że proporcja ta jest prawdziwa również wtedy, gdy $a_x = a_y = 0$ dla dowolnych b_x, b_y oraz gdy $a_x = 0$, $a_y \neq 0$ dla $b_x = 0$ i dowolnego b_y. Wówczas otrzymujemy łatwy do zapamiętania warunek (1') równoległości dowolnych wektorów.

Na przykład wektory [0,3,6], [0,1,2] są równoległe, gdyż spełniają warunek (1') równoległości wektorów (przy tej dodatkowej umowie).

P R Z Y K Ł A D 1.3. Znajdziemy wektor prostopadły do wektorów $\vec{a} = [4,2,3]$ i $\vec{b} = [2,6,4]$.

Wektorem prostopadłym do wektorów $\vec{a}$ i $\vec{b}$ jest na przykład iloczyn wektorowy tych wektorów:

$$\vec{w} = \vec{a}\times\vec{b} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 4 & 2 & 3 \\ 2 & 6 & 4 \end{vmatrix} = -10\vec{i} - 10\vec{j} + 20\vec{k} = [-10,-10,20].$$

Ponieważ wektory [10,10,−20], [−1,−1,2], [−2,−2,4], [1,1,−2] są równoległe do wektora $\vec{w}$, więc są one również prostopadłe do wektorów $\vec{a}$ i $\vec{b}$. ■

P R Z Y K Ł A D 1.4. Dobierzemy parametr m tak, by wektory $[m-2,2,5]$, $[m^2-12,m,10]$ były równoległe.

Wektory te są równoległe wtedy i tylko wtedy, gdy ich współrzędne są proporcjonalne. Ponieważ

$$\frac{m-2}{m^2-12} = \frac{2}{m} = \frac{5}{10} \Leftrightarrow \left(\frac{m-2}{m^2-12} = \frac{5}{10} \wedge \frac{2}{m} = \frac{5}{10}\right) \Leftrightarrow$$

$$\Leftrightarrow [(m=4 \vee m=-2) \wedge m=4],$$

więc wektory te są równoległe dla $m = 4$. ■

P R Z Y K Ł A D 1.5. Obliczmy pole S trójkąta o wierzchołkach A(3,2,-1), B(2,4,0), C(5,7,0).

Pole S trójkąta „zbudowanego" na wektorach

$$\overrightarrow{AB} = [-1,2,1], \quad \overrightarrow{AC} = [2,5,1]$$

jest równe połowie pola równoległoboku „zbudowanego" na tych wektorach; pole S jest więc równe połowie długości iloczynu wektorowego tych wektorów. Mamy więc

$$S = \frac{1}{2}\left|\overrightarrow{AB}\times\overrightarrow{AC}\right| = \frac{3}{2}\sqrt{11},$$

gdyż

$$\overrightarrow{AB}\times\overrightarrow{AC} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -1 & 2 & 1 \\ 2 & 5 & 1 \end{vmatrix} = -3\vec{i} + 3\vec{j} - 9\vec{k}, \quad \left|\overrightarrow{AB}\times\overrightarrow{AC}\right| = \sqrt{9+9+81} = 3\sqrt{11}. \blacksquare$$

P R Z Y K Ł A D 1.6. Pokażemy, że punkty: P(1,2,1), Q(3,4,5), R(4,5,7) leżą na jednej prostej.

W tym celu wystarczy pokazać, że $\overrightarrow{PQ} \parallel \overrightarrow{PR}$. Obliczymy współrzędne tych wektorów:

$$\overrightarrow{PQ} = [2,2,4], \quad \overrightarrow{PR} = [3,3,6].$$

Ponieważ współrzędne tych wektorów są proporcjonalne:

$$\frac{2}{3} = \frac{2}{3} = \frac{4}{6},$$

więc wektory $\overrightarrow{PQ}$ i $\overrightarrow{PR}$ są równoległe. Punkty P, Q, R leżą na jednej prostej ■

ILOCZYN MIESZANY. Iloczynem mieszanym uporządkowanej trójki wektorów $\vec{a},\vec{b},\vec{c}$ w układzie 0xyz nazywamy liczbę $(\vec{a},\vec{b},\vec{c})$ określoną wzorem

$$(\vec{a},\vec{b},\vec{c}) \overset{\text{def}}{=} (\vec{a}\times\vec{b})\circ\vec{c}.$$

Iloczyn mieszany jest liczbą, gdyż $\vec{a}\times\vec{b}$ jest wektorem, zaś iloczyn skalarny wektorów $\vec{a}\times\vec{b}$ oraz $\vec{c}$ jest liczbą.

Z definicji iloczynu mieszanego wynikają następujące własności:

(1) $(\vec{a}\times\vec{b})\circ\vec{c}=-(\vec{b}\times\vec{a})\circ\vec{c}$,

(2) $(\vec{a}\times\vec{b})\circ\vec{c}=\vec{a}\circ(\vec{b}\times\vec{c})$,

(3) $(\vec{a}\times\vec{b})\circ\vec{c}=(\vec{b}\times\vec{c})\circ\vec{a}=(\vec{c}\times\vec{a})\circ\vec{b}$ przemienność cykliczna.

Z własności (3) wynika na przykład, że

$$(\vec{i}\times\vec{j})\circ\vec{k}=(\vec{j}\times\vec{k})\circ\vec{i}=(\vec{k}\times\vec{i})\circ\vec{j}=\vec{j}\circ\vec{j}=1.$$

TWIERDZENIE 1.5 (podstawowe o iloczynie mieszanym). Jeżeli

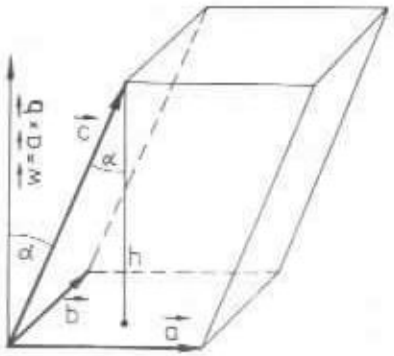
$$\vec{a}=[a_x,a_y,a_z],\quad \vec{b}=[b_x,b_y,b_z],\quad \vec{c}=[c_x,c_y,c_z],$$

to iloczyn mieszany $(\vec{a},\vec{b},\vec{c})$ wyraża się wzorem

$$(\vec{a},\vec{b},\vec{c})=\begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}. \tag{1.4}$$

Iloczyn mieszany ma następującą interpretację geometryczną:

wartość bezwzględna iloczynu mieszanego niezerowych i wzajemnie nierównoległych wektorów $\vec{a},\vec{b},\vec{c}$ jest równa objętości $|V|$ równoległościanu "zbudowanego" na tych wektorach. Istotnie objętość $|V|$ jest równa iloczynowi pola S równoległoboku zbudowanego na wektorach $\vec{a}$, $\vec{b}$ i wysokości h, czyli



$$|V|=S\,h=|\vec{a}\times\vec{b}|\,h=|\vec{a}\times\vec{b}||\vec{c}||\cos\alpha|=$$
$$=||\vec{a}\times\vec{b}||\vec{c}|\cos\alpha\,|=$$
$$=||\vec{a}\times\vec{b}||\vec{c}|\cos\sphericalangle(\vec{a}\times\vec{b},\vec{c})|=|(\vec{a}\times\vec{b})\circ\vec{c}|.$$

PRZYKŁAD 1.7. Obliczymy objętość $|V|$ ostrosłupa o wierzchołkach: A(1,1,1), B(2,3,4), C(-1,4,1), D(5,5,5).

Rozważmy równoległościan i ostrosłup „zbudowany" na wektorach

$$\overrightarrow{AB}=[1,2,3],\qquad \overrightarrow{AC}=[-2,3,0],\qquad \overrightarrow{AD}=[4,4,4].$$

Objętość $|V_1|$ tego równoległościanu jest równa iloczynowi pola S równoległoboku „zbudowanego" na wektorach $\overrightarrow{AB}$, $\overrightarrow{AC}$ i wysokości h i jest równa wartości bezwzględnej iloczynu mieszanego wektorów $\overrightarrow{AB}$, $\overrightarrow{AC}$, $\overrightarrow{AD}$:

$$|V_1|=S\cdot h=\left|\begin{vmatrix} 1 & 2 & 3 \\ -2 & 3 & 0 \\ 4 & 4 & 4 \end{vmatrix}\right|=|-32|=32.$$

Natomiast objętość $|V|$ rozważanego ostrosłupa jest równa 1/3 iloczynu pola trójkąta zbudowanego na wektorach $\overrightarrow{AB}$, $\overrightarrow{AC}$ i wysokości h (pole tego trójkąta jest równe 1/2 pola S). Zatem

$$|V|=\frac{1}{3}\cdot\frac{1}{2}S\cdot h=\frac{1}{6}\cdot S\cdot h=\frac{1}{6}|V_1|=\frac{32}{6}=\frac{16}{3}.$$ ■

PRZYKŁAD 1.8. Sprawdzimy, że punkty: A(1,2,1), B(3,4,5), C(3,5,7), D(2,4,5) leżą w jednej płaszczyźnie.

Interpretacja geometryczna iloczynu mieszanego pozwala nam wnioskować następująco:

– jeżeli iloczyn mieszany wektorów $\overrightarrow{AB}$, $\overrightarrow{AC}$, $\overrightarrow{AD}$ jest różny od zera, to punkty A, B, C, D nie leżą w jednej płaszczyźnie,

– jeżeli iloczyn mieszany wektorów $\overrightarrow{AB}$, $\overrightarrow{AC}$, $\overrightarrow{AD}$ jest równy zeru, to punkty A, B, C, D leżą w jednej płaszczyźnie.

Ponieważ

$$\overrightarrow{AB}=[2,2,4],\quad \overrightarrow{AC}=[2,3,6],\quad \overrightarrow{AD}=[1,2,4],$$

więc

$$(\overrightarrow{AB}\times\overrightarrow{AC})\circ\overrightarrow{AD}=\begin{vmatrix} 2 & 2 & 4 \\ 2 & 3 & 6 \\ 1 & 2 & 4 \end{vmatrix}=0.$$

Oznacza to, że punkty A, B, C, D leżą w jednej płaszczyźnie. ■

ZADANIA DO ROZWIĄZANIA.

1. Obliczyć iloczyn skalarny wektorów:
   a) [3,4], [2,-3], b) [3,4,1], [2,-3,0], c) $\bar{u}=2\bar{i}-\bar{j}$, $\bar{v}=\bar{i}+4\bar{k}$.

2. Obliczyć pracę, jaką wykona siła $\bar{F}=[3,4,1]$ przesuwając punkt materialny wzdłuż:
   a) osi 0x od punktu o odciętej $x=4$ do punktu o odciętej $x=-3$,
   b) linii prostej od punktu P(3,3,8) do punktu Q(5,6,0).

3. Obliczyć iloczyn wektorowy wektorów z zadania 1.

4. Sprawdzić, czy wektory $\bar{u}$ i $\bar{v}$ są równoległe, czy prostopadłe, jeśli:
   a) $\bar{u}=[-1,0,3]$, $\bar{v}=[3,0,-9]$, b) $\bar{u}=[-1,0,3]$, $\bar{v}=[6,7,2]$.

5. Czy można dobrać parametr k tak, by wektory $\bar{u}$ i $\bar{v}$ były pro stopadłe, jeśli:
   a) $\bar{u}=[k,0,1]$, $\bar{v}=[-1,k,k]$ b) $\bar{u}=[k,k,k+4]$, $\bar{v}=[k+1,3,9]$,
   c) $\bar{u}=[2,k,1]$, $\bar{v}=[k,2k,2]$, d) $\bar{u}=[k,3,4]$, $\bar{v}=[1,k,1]$?

6. Czy można dobrać parametr k tak, by wektory $\bar{u}$ i $\bar{v}$ z zadania 5 były równoległe?

7. Znaleźć trzy wektory równoległe do wektora [4,2,-8].

8. Znaleźć trzy wektory prostopadłe do wektora [4,2,-8].

9. Obliczyć cosinus i sinus kąta φ między wektorami $\bar{u}$ i $\bar{v}$, jeśli:
   a) $\bar{u}=[1,-2,2]$, $\bar{v}=[2,1,-2]$, b) $\bar{u}=[0,3,-4]$, $\bar{v}=[2,2,1]$,
   c) $\bar{u}=[1,1,1]$, $\bar{v}=[5,1,1]$, d) $\bar{u}=[5,0,-3]$, $\bar{v}=[0,4,0]$.

10. Znaleźć współrzędną prostokątnego rzutu wektora: $\bar{u}$ na oś wektora $\bar{v}$, $\bar{v}$ na oś wektora $\bar{u}$, $\bar{u}$ na oś 0x, wersora $\bar{i}$ na oś wektora $\bar{u}$, gdy
    a) $\bar{u}=[-2,-2,1]$, $\bar{v}=[0,-6,8]$, b) $\bar{u}=[4,2,4]$, $\bar{v}=[1,2,2]$.

11. Obliczyć pole trójkąta o wierzchołkach:
    a) A(1,2,3), B(-1,0,4), C(5,6,0),
    b) A(1,2,0), B(-1,0,0), C(5,-6,0),
    c) A(0,0,0), B(3,4,5), C(0,0,6).

12. Obliczyć pole równoległoboku ABCD i znaleźć punkt D, jeśli:
    a) A(1,2,3), B(4,0,3), C(-2,3,0),
    b) A(0,0,0), B(5,0,3), C(1,1,1),
    c) A(-1,2,3), B(4,5,6), C(0,-1,2).

13. Sprawdzić, czy punkty P, Q, R leżą na jednej prostej, jeśli:
    a) P(0,0,3), Q(-1,2,4), R(2,-4,1),
    b) P(1,2,-1), Q(3,0,2), R(-1,1,1),
    c) P(-1,0,0), Q(5,6,7), R(-13,-12,-14).

14. Sprawdzić, czy punkty P, Q, R, S leżą na jednej płaszczyźnie, jeśli:
    a) P(0,3,4), Q(−1,2,2), R(2,0,3), S(-1,1,1),
    b) P(1,1,-13), Q(−1,0,−14), R(0,-4,0), S(3,-2,0),
    c) P(-3,2,-2), Q(-1,1,-2), R(-3,4,1), S(-2,1,0).

Odpowiedzi.

1. a) -6, b) -6, c) 2.
2. a) -21, b) 10.
3. a) [0,0,-17], b) [3,2,-17], c) [-4,-8,1].
4. a) Równoległe, b) prostopadłe.
5. a) Dla dowolnego $k\in R$, b) $k=-4$ lub $k=-9$, c) nie, d) $k=-1$.
6. a) Nie, b) $k=2$, c) $k=4$, d) nie.
7. Wsk. Każdy wektor [4s,2s,-8s], $s\in R-\{0\}$ jest wektorem równoległym do [4,2,-8].
8. Wsk. $[a_x,a_y,a_z]\perp[4,2,-8]\Leftrightarrow 4a_x+2a_y-8a_z=0$.
9. a) $\cos\varphi=-4/9$, $\sin\varphi=\sqrt{65}/9$, b) $\cos\varphi=2/15$, $\sin\varphi=\sqrt{221}/15$,
   c) $\cos\varphi=7/9$, $\sin\varphi=4\sqrt{2}/9$, d) $\cos\varphi=0$, $\sin\varphi=1$.
10. a) 2, 20/3, -2, -2/3, b) 16/3, 8/3, 4, 2/3.
11. a) $\sqrt{2}$, b) 12, c) 15.
12. a) $3\sqrt{14}$, D(-5,5,0), b) $\sqrt{38}$, D(-4,1,-2), c) $2\sqrt{106}$, D(-5,-4,-1).
13. a) Tak, b) nie, c) tak.
14. a) Tak, b) tak, c) nie.

2. PŁASZCZYZNA.

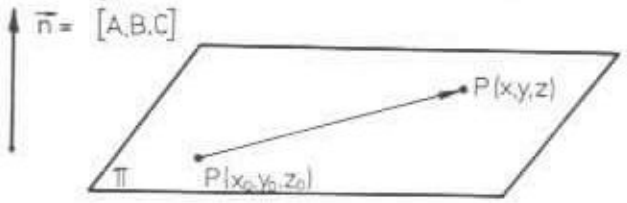
RÓWNANIE OGÓLNE PŁASZCZYZNY. Płaszczyznę w przestrzeni trójwymiarowej wyznaczają m.inn. punkt leżący na tej płaszczyźnie i wektor prostopadły do tej płaszczyzny.

TWIERDZENIE 2.1. Załóżmy, że niezerowy wektor $\vec{n}=[A,B,C]$ (tzn. $A^2+B^2+C^2>0$) jest prostopadły do płaszczyzny π i punkt $P_0(x_0,y_0,z_0)$ należy do tej płaszczyzny (rys 2.1).

Wówczas punkt $P(x,y,z)$ należy do płaszczyzny π wtedy i tylko wtedy, gdy liczby x, y, z spełniają równanie liniowe

(2.1) $$A(x-x_0)+B(y-y_0)+C(z-z_0)=0.$$

D o w ó d .



Rys 2.1.

Niech $P(x,y,z)$ będzie dowolnym, różnym od P_0, punktem w układzie współrzędnych $0xyz$. Wówczas

$$P(x,y,z)\in\pi \Leftrightarrow \overrightarrow{P_0P}\parallel\pi \Leftrightarrow \overrightarrow{P_0P}\perp\vec{n} \Leftrightarrow \vec{n}\circ\overrightarrow{P_0P}=0 \Leftrightarrow$$
$$\Leftrightarrow [A,B,C]\circ[x-x_0,y-y_0,z-z_0]=0 \Leftrightarrow$$
$$\Leftrightarrow A(x-x_0)+B(y-y_0)+C(z-z_0)=0.$$

Łatwo sprawdzić, że współrzędne (x_0,y_0,z_0) punktu P_0 też spełniają to równanie. □

Jeżeli w równaniu (2.1) wykonamy działania i oznaczymy $-Ax_0-By_0-Cz_0\equiv D$, to otrzymamy równanie płaszczyzny w postaci

(2.2) $$Ax+By+Cz+D=0, \quad \text{gdzie } A^2+B^2+C^2>0,$$

zwane **równaniem ogólnym** płaszczyzny.

Na przykład płaszczyzna π przechodząca przez punkt $P_0(3,1,-1)$ i prostopadła do wektora $[2,0,-3]$ ma równanie

$$\pi: 2(x-3)+0(y-1)-3(z+1)=0,$$

czyli

$$\pi: 2x-3z-9=0.$$

P R Z Y K Ł A D 2.1. Napiszemy równanie płaszczyzny π przechodzącej przez punkt $P_0(2,3,-1)$ i równoległej do wektorów $\vec{r}_1=[1,2,3]$, $\vec{r}_2=[-1,3,0]$.

Punkt $P_0(2,3,-1)$ należy do płaszczyzny π, zatem

$$\pi: A(x-2)+B(y-3)+C(z+1)=0.$$

Płaszczyzna π jest równoległa do wektorów $\vec{r}_1$, $\vec{r}_2$, zatem wektor $\vec{n}=[A,B,C]$ jest do tych wektorów prostopadły. Możemy więc przyjąć

$$\vec{n}=[A,B,C]=\vec{r}_1\times\vec{r}_2=[1,2,3]\times[-1,3,0]=[-9,-3,5].$$

Płaszczyzna π ma więc równanie

$$\pi: -9(x-2)-3(y-3)+5(z+1)=0,$$

czyli

$$\pi: 9x+3y-5z-32=0$$ ■

P R Z Y K Ł A D 2.2. (równanie płaszczyzny przechodzącej przez trzy punkty). Napiszemy równanie płaszczyzny π przechodzącej przez punkty $P(1,2,3)$, $Q(0,3,5)$, $R(-1,2,0)$, (rys 2.2).

Punkt $P(1,2,3)$ należy do płaszczyzny π, zatem

$$\pi: A(x-1)+B(y-2)+C(z-3)=0.$$

Wektor $\vec{n}=[A,B,C]$ prostopadły do płaszczyzny π spełnia warunki

$$\vec{n}\perp\overrightarrow{PQ} \quad \text{oraz} \quad \vec{n}\perp\overrightarrow{PR},$$

a zatem możemy przyjąć

$$\vec{n}=[A,B,C]=\overrightarrow{PQ}\times\overrightarrow{PR}=[-1,1,2]\times[-2,0,-3]=[-3,-7,2].$$

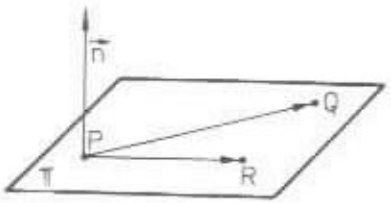
Płaszczyzna π ma więc równanie

$$\pi: -3(x-1)-7(y-2)+2(z-3)=0$$

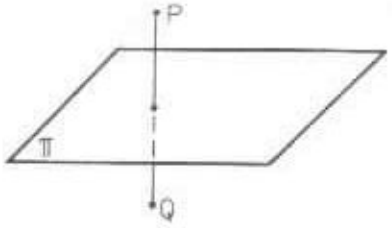
czyli

$$\pi: 3x+7y-2z-11=0.$$ ■

U w a g a . Równanie płaszczyzny π w tym przykładzie możemy również znaleźć przyjmując np. $(x_0,y_0,z_0)=(0,3,5)$, $[A,B,C]=\overrightarrow{QR}\times\overrightarrow{QP}$.



Rys 2.2.



Rys 2.3.

P R Z Y K Ł A D 2.3. Napiszemy równanie płaszczyzny π przechodzącej przez środek odcinka $\overline{PQ}$ i prostopadłej do tego odcinka, gdy P(3,4,0), Q(−7,8,6), (rys. 2.3).

Płaszczyzna π przechodzi przez środek odcinka $\overline{PQ}$, czyli przez punkt

$$(\frac{3-7}{2},\frac{4+8}{2},\frac{0+6}{2})=(-2,6,3),$$

zatem

$$\pi\colon A(x+2)+B(y-6)+C(z-3)=0.$$

Wektorem $\vec{n}=[A,B,C]$ prostopadłym do płaszczyzny jest np. wektor $\overrightarrow{PQ}$, czyli

$$[A,B,C]=\overrightarrow{PQ}=[-10,4,6],$$

a stąd $$\pi\colon -10(x+2)+4(y-6)+6(z-3)=0,$$

czyli $$\pi\colon 5x-2y-3z+31=0.$$ ■

SZCZEGÓLNE POŁOŻENIA PŁASZCZYZNY W UKŁADZIE WSPÓŁRZĘDNYCH 0xyz. Rozpatrzmy szczególne położenie płaszczyzny $\pi\colon Ax+By+Cz+D=0$ w układzie współrzędnych 0xyz.

1) Płaszczyzna $\pi\colon By+Cz+D=0$ (mamy tu $A=0$) jest równoległa do osi 0x, gdyż

$$A=0 \Leftrightarrow \vec{n}=[0,B,C] \Leftrightarrow \vec{n}\perp 0x \Leftrightarrow \pi\parallel 0x,$$

Analogicznie: płaszczyzna dana równaniem $Ax+Cz+D=0$ jest równoległa do osi 0y, zaś płaszczyzna dana równaniem $Ax+By+D=0$ jest równoległa do osi 0z.

2) Płaszczyzna $\pi\colon Cz+D=0$ (mamy tu $A=0$, $B=0$) jest równoległa do płaszczyzny 0xy, czyli jest prostopadła do osi 0z, gdyż

$$A=B=0 \Leftrightarrow \vec{n}=[0,0,C] \Leftrightarrow \vec{n}\perp 0xy \Leftrightarrow \pi\parallel 0xy \Leftrightarrow \pi\perp 0z.$$

Analogicznie: płaszczyzna dana równaniem $Ax+D=0$ jest równoległa do płaszczyzny 0yz, czyli jest prostopadła do osi 0x, zaś płaszczyzna dana równaniem $By+D=0$ jest równoległa do płaszczyzny 0xz, czyli jest prostopadła do osi 0y.

3) Płaszczyzna dana równaniem $Ax+By+Cz=0$ (mamy tu $D=0$) przechodzi przez początek układu (0,0,0).

Płaszczyzna dana równaniem $Ax+By=0$ zawiera oś Oz, gdyż jest równoległa do osi Oz i przechodzi przez punkt (0,0,0).

W szczególności płaszczyzna Oxy ma równanie $z=0$, płaszczyzna Oxz ma równanie $y=0$, zaś płaszczyzna Oyz ma równanie $x=0$.

P R Z Y K Ł A D 2.4. Napiszemy równanie płaszczyzny π przechodzącej przez punkt P(1,2,3) i zawierającej oś Ox.

1 s p o s ó b . Płaszczyzna π przechodzi przez punkt P(1,2,3), zatem

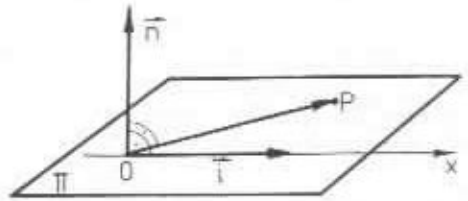
$$\pi\colon A(x-1)+B(y-2)+C(z-3)=0.$$

Pozostaje znaleźć dowolny wektor $\vec{n}=[A,B,C]$ prostopadły do płaszczyzny π. Z warunków zadania wynika, że

$$\overrightarrow{0P}=[1,2,3]\parallel\pi, \qquad \vec{i}=[1,0,0]\parallel\pi,$$

zatem (rys 2.4) wektor $\vec{n}=\overrightarrow{0P}\times\vec{i}$ jest prostopadły do płaszczyzny π:

$$\vec{n}=[A,B,C]=\overrightarrow{0P}\times\vec{i}=[1,2,3]\times[1,0,0]=[0,3,-2]\perp\pi.$$



Rys 2.4.

Płaszczyzna π ma więc równanie

$$\pi\colon 0(x-1)+3(y-2)-2(z-3)=0,$$

czyli

$$\pi: 3y-2z=0.$$

II sposób. Z warunków zadania wynika, że

$$(0,0,0)\in\pi \quad \text{i} \quad 0x\parallel\pi,$$

zatem $D=0$ oraz $A=0$, a stąd

$$\pi: By+Cz=0.$$

Punkt P(1,2,3) należy do płaszczyzny π, co równoważne jest warunkowi $2B+3C=0$. Równanie to ma oczywiście nieskończenie wiele rozwiązań. Przyjmując np. $C=-4$ otrzymujemy $B=6$. Szukana płaszczyzna ma równanie

$$\pi: 6y-4z=0$$

równoważne równaniu

$$\pi: 3y-2z=0. \qquad \blacksquare$$

WZAJEMNE POŁOŻENIE DWU PŁASZCZYZN. Położenie dwu płaszczyzn będziemy rozpatrywać ze względu na kąt dwuścienny między tymi płaszczyznami.

Przypomnijmy, że kąt liniowy kąta dwuściennego między dwiema płaszczyznami ma miarę z przedziału $<0,\frac{\pi}{2}>$.

TWIERDZENIE 2.2. Załóżmy, że płaszczyzny π_1, π_2 mają równania

$$\pi_1: A_1x+B_1y+C_1z+D_1=0,$$
$$\pi_2: A_2x+B_2y+C_2z+D_2=0.$$

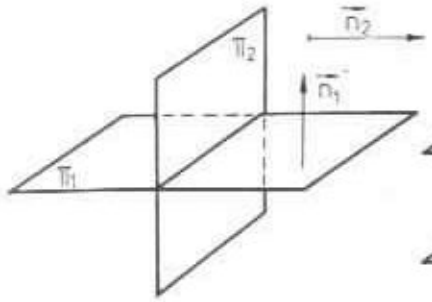
Wówczas

(1) $\pi_1\perp\pi_2 \Leftrightarrow A_1A_2+B_1B_2+C_1C_2=0$, (rys 2.5),

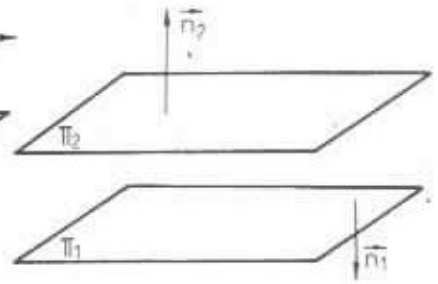
(2) $\pi_1\parallel\pi_2 \Leftrightarrow \frac{A_1}{A_2}=\frac{B_1}{B_2}=\frac{C_1}{C_2}$, (rys 2.6),

(3) $$\cos\sphericalangle(\pi_1,\pi_2)=\frac{|A_1A_2+B_1B_2+C_1C_2|}{\sqrt{A_1^2+B_1^2+C_1^2}\cdot\sqrt{A_2^2+B_2^2+C_2^2}}.$$

Dowód (3). Niech α będzie kątem między wektorami prostopadłymi do płaszczyzn π_1, π_2, zaś φ kątem między tymi płaszczyznami (rys 2.7).



Rys 2.5.



Rys 2.6.

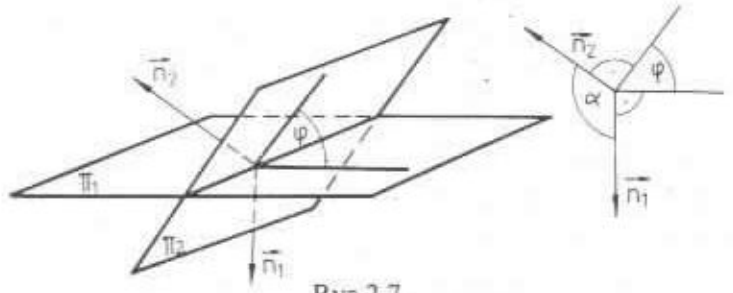
Można zauważyć (rys 2.7), że

$$\varphi=\begin{cases}\alpha, & \text{gdy } \alpha\in<0,\pi/2>\\ \pi-\alpha, & \text{gdy } \alpha\in(\pi/2,\pi>\end{cases}$$

Wobec tego

$$\cos\varphi=|\cos\alpha|.$$

Uwzględniając następnie wzór na cosinus kąta α między wektorami $\bar{n}_1=[A_1,B_1,C_1]$ oraz $\bar{n}_2=[A_2,B_2,C_2]$ otrzymujemy tezę (3). $\square$



Rys 2.7.

PRZYKŁAD 2.5. Napiszemy równanie płaszczyzny π przechodzącej przez punkt $P_0(1,2,3)$ oraz prostopadłej do płaszczyzn

$$\pi_1: 2x-3y+z-4=0, \qquad \pi_2: x-y-3z+10=0.$$

I sposób. Punkt $P_0(1,2,3)$ należy do płaszczyzny π, zatem

$$\pi: A(x-1)+B(y-2)+C(z-3)=0.$$

Z warunków zadania mamy: $\pi\perp\pi_1$, $\pi\perp\pi_2$. Ponieważ

$$\pi \perp \pi_1 \Leftrightarrow [A,B,C] \perp [2,-3,1],$$
$$\pi \perp \pi_2 \Leftrightarrow [A,B,C] \perp [1,-1,-3],$$

więc możemy przyjąć

$$[A,B,C] = [2,-3,1] \times [1,-1,-3] = [10,7,1].$$

Zatem płaszczyzna π ma równanie

$$\pi: 10(x-1)+7(y-2)+1(z-3)=0,$$

czyli

$$\pi: 10x+7y+z-27=0.$$

II sposób.

$$\left.\begin{array}{lll} P(1,2,3)\in\pi & \Leftrightarrow & \pi: A(x-1)+B(y-2)+C(z-3)=0, \\ \pi\perp\pi_1 & \Leftrightarrow & A\cdot 2+B\cdot(-3)+C\cdot 1=0, \\ \pi\perp\pi_2 & \Leftrightarrow & A\cdot 1+B\cdot(-1)+C\cdot(-3)=0, \end{array}\right\}.$$

Jest to układ równań liniowych jednorodnych z niewiadomymi A, B, C. Z założenia, że wektor [A,B,C] jest wektorem niezerowym wynika, że układ ten ma rozwiązanie niezerowe. Zatem, zgodnie z twierdzeniem Cramera, wyznacznik główny tego układu jest równy zeru:

$$\begin{vmatrix} x-1 & y-2 & z-3 \\ 2 & -3 & 1 \\ 1 & -1 & -3 \end{vmatrix} = 0.$$

Jest to równanie szukanej płaszczyzny. Po obliczeniu wyznacznika i uporządkowaniu otrzymujemy:

$$\pi: 10x+7y+z-27=0 \qquad \blacksquare$$

PRZYKŁAD 2.6. Napiszemy równanie płaszczyzny π przechodzącej przez punkty P(1,0,4) i Q(4,3,−2) oraz prostopadłej do płaszczyzny π_1: $2x-y-3z+7=0$.

I sposób. Punkt P(1,0,4) należy do płaszczyzny π, zatem

$$\pi: A(x-1)+B(y-0)+C(z-4)=0.$$

Z warunków zadania wynika, że $[A,B,C] \perp \overrightarrow{PQ}$ i $\pi \perp \pi_1$. Ponieważ

$$[A,B,C] \perp \overrightarrow{PQ} \Leftrightarrow [A,B,C] \perp [3,3,-6],$$
$$\pi \perp \pi_1 \Leftrightarrow [A,B,C] \perp [2,-1,-3],$$

więc możemy przyjąć, że

$$[A,B,C] = [3,3,-6] \times [2,-1,-3] = [-15,-3,-9].$$

Płaszczyzna π ma więc równanie

$$\pi: -15(x-1)-3y-9(z-4)=0,$$

czyli

$$\pi: 5x+y+3z-17=0.$$

II sposób. Z warunków zadania mamy

$$\left.\begin{array}{lllll} P(1,0,4)\in\pi & \Leftrightarrow & \pi: A(x-1)+B(y-0)+C(z-4)=0 \\ Q(4,3,-2)\in\pi & \Leftrightarrow & 3A & +3B & -6C & =0 \\ \pi\perp\pi_1 & \Leftrightarrow & 2A & -B & -3C & =0 \end{array}\right\}.$$

Stąd, analogicznie jak w przykładzie 2.5, otrzymujemy

$$\begin{vmatrix} x-1 & y & z-4 \\ 3 & 3 & -6 \\ 2 & -1 & -3 \end{vmatrix} = 0,$$

czyli $5x+y+3z-17=0$. Jest to równanie szukanej płaszczyzny π. $\blacksquare$

ODLEGŁOŚĆ PUNKTU OD PŁASZCZYZNY. Niech płaszczyzna π ma równanie

$$\pi: Ax+By+Cz+D=0.$$

Wykazuje się, że odległość d punktu (x_0,y_0,z_0) od płaszczyzny π wyraża się wzorem

$$d = \frac{|Ax_0+By_0+Cz_0+D|}{\sqrt{A^2+B^2+C^2}}. \tag{2.3}$$

Jest to wzór analogiczny do wzoru

$$d = \frac{|Ax_0+By_0+C|}{\sqrt{A^2+B^2}}$$

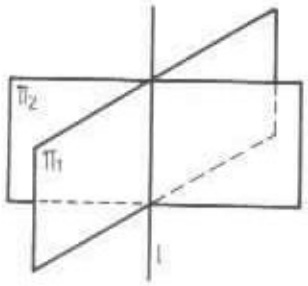
na odległość punktu (x_0,y_0) od prostej $Ax+By+C=0$ na płaszczyźnie $0xy$.

Dla przykładu odległość punktu (2,1,3) od płaszczyzny $\pi: 2x-y-2z-3=0$ jest równa

$$d = \frac{|2\cdot 2-1\cdot 1-2\cdot 3-3|}{\sqrt{2^2+(-1)^2+(-2)^2}} = 2.$$

W ten sam sposób otrzymujemy,że odległość punktu (2,-3,2) od tej płaszczyzny jest równa zeru, co oznacza, że punkt (2,-3,2) należy do płaszczyzny π.

PĘK PŁASZCZYZN. Niech l oznacza krawędź przecięcia dwu nierównoległych płaszczyzn π_1 i π_2 (rys 2.8).



Rys 2.8.

Pękiem płaszczyzn wyznaczonym przez dwie nierównoległe płaszczyzny π_1 i π_2 nazywamy zbiór wszystkich płaszczyzn zawierających krawędź l przecięcia płaszczyzn π_1 i π_2.

Oczywiście do tego pęku płaszczyzn należą również płaszczyzny π_1 i π_2.

TWIERDZENIE 2.3. Załóżmy, że dwie nierównoległe płaszczyzny mają równania

$$\pi_1: A_1x+B_1y+C_1z+D_1=0,$$
$$\pi_2: A_2x+B_2y+C_2z+D_2=0.$$

Wówczas

(1) dowolna płaszczyzna π należąca do pęku płaszczyzn wyznaczonego przez płaszczyzny π_1 i π_2 ma równanie postaci

$$(2.4)\quad \pi: \lambda_1(A_1x+B_1y+C_1z+D_1)+\lambda_2(A_2x+B_2y+C_2z+D_2)=0,$$

gdzie λ_1,λ_2 są pewnymi liczbami rzeczywistymi nierównymi jednocześnie zeru tzn. $\lambda_1,\lambda_2\in\mathbb{R}$, $\lambda_1^2+\lambda_2^2>0$,

(2) dla dowolnych $\lambda_1,\lambda_2\in\mathbb{R}$, $\lambda_1^2+\lambda_2^2>0$, płaszczyzna π dana równaniem (2.4) należy do pęku wyznaczonego przez płaszczyzny π_1 i π_2.

Zauważmy, że dla $\lambda_1=0$ z (2.4) otrzymujemy równanie płaszczyzny π_2, zaś dla $\lambda_2=0$ - równanie płaszczyzny π_1.

Równanie (2.4) nazywamy **równaniem pęku** płaszczyzn wyzmaczonego przez płaszczyzny π_1 i π_2.

PRZYKŁAD 2.7. Napiszemy równanie płaszczyzny π przechodzącej przez punkt (1,3,4) oraz zawierającej krawędź przecięcia płaszczyzn π_1: $x-y-3z-6=0$, π_2: $x+y+2z+3=0$.

Szukana płaszczyzna π jest jedną z pęku płaszczyzn wyznaczonego przez płaszczyzny π_1 i π_2. Zatem

$$\pi: \lambda_1(x-y-3z-6)+\lambda_2(x+y+2z+3)=0.$$

Ponieważ punkt (1,3,4) należy do płaszczyzny π, więc

$$\lambda_1(1-3-3\cdot4-6)+\lambda_2(1+3+2\cdot4+3)=0.$$

Stąd otrzymujemy związek między λ_1 i λ_2:

$$-20\lambda_1+15\lambda_2=0.$$

Zatem np. dla $\lambda_1=3$ mamy $\lambda_2=4$. Płaszczyzna π ma więc równanie

$$\pi: 3(x-y-3z-6)+4(x+y+2z+3)=0,$$

czyli

$$\pi: 7x+y-z-6=0 \qquad \blacksquare$$

PRZYKŁAD 2.8. Napiszemy równanie płaszczyzny π zawierającej krawędź przecięcia płaszczyzn

$$\pi_1: x+y-2z+3=0, \qquad \pi_2: 2x-y-3z-6=0$$

oraz prostopadłej do płaszczyzny

$$\pi_3: 3x-y+4z-7=0.$$

Szukana płaszczyzna π jest jedną z pęku płaszczyzn wyznaczonego przez płaszczyzny π_1 i π_2, zatem

$$\pi: \lambda_1(x+y-2z+3)+\lambda_2(2x-y-3z-6)=0,$$

czyli

$$\pi: (\lambda_1+2\lambda_2)x+(\lambda_1-\lambda_2)y-(2\lambda_1+3\lambda_2)z+3\lambda_1-6\lambda_2=0.$$

Z warunków zadania wiemy, że $\pi\perp\pi_3$. Ponieważ

$$\pi\perp\pi_3 \Leftrightarrow (\lambda_1+2\lambda_2)\cdot3+(\lambda_1-\lambda_2)(-1)-(2\lambda_1+3\lambda_2)\cdot4=0,$$

więc otrzymujemy

$$-6\lambda_1-5\lambda_2=0.$$

Zatem np. dla $\lambda_1=5$, mamy $\lambda_2=-6$. Płaszczyzna π ma więc równanie

$$\pi: 5(x+y-2z+3)-6(2x-y-3z-6)=0,$$

czyli $\pi: 7x-11y-8z-51=0.$ $\blacksquare$

ZADANIA DO ROZWIĄZANIA.

1. a) Jakie elementy wyznaczają płaszczyznę?
   b) Czy dwa różne punkty wyznaczają płaszczyznę?
   c) Czy punkt leżący na płaszczyźnie i wektor równoległy do płaszczyzny wyznaczają płaszczyznę?

2. Napisać równanie płaszczyzny π przechodzącej przez punkt P_0 i równoległej do płaszczyzny π_1, gdy:
   a) $P_0=(3,-2,1)$, π_1: $2x-2y-4z-7=0$,
   b) $P_0=(0,0,0)$, π_1: $x+3z-11=0$,
   c) $P_0=(2,3,0)$, π_1 jest płaszczyzną 0xy,
   d) $P_0=(2,3,0)$, π_1 jest płaszczyzna 0xz.

3. Napisać równanie płaszczyzny π przechodzącej przez punkty P_1, P_2 i prostopadłej do płaszczyzny π_1, gdy:
   a) $P_1=(6,2,1)$, $P_2=(3,-1,1)$, π_1: $x+2y-z-6=0$,
   b) $P_1=(-2,0,3)$, $P_2=(1,1,1)$, π_1: $2x-3z-8=0$,
   c) $P_1=(1,2,4)$, $P_2=(-2,4,5)$, π_1 jest płaszczyzną 0xy.

4. Napisać równanie płaszczyzny π przechodzącej przez punkt P_0 i prostopadłej do płaszczyzn π_1 i π_2, gdy:
   a) $P_0=(3,-2,1)$, π_1: $2x-2y-4z-7=0$, π_2: $x+y-3z-1=0$,
   b) $P_0=(0,0,0)$, π_1: $x+3z-11=0$, π_2: $x+2y-3z=0$,
   c) $P_0=(1,3,4)$, π_1: $x-z-3=0$, π_2 jest płaszczyzną 0xy,
   d) $P_0=(1,3,4)$, π_1 jest płaszczyzną 0xy, π_2 - płaszczyzną 0xz.

5. Napisać równanie płaszczyzny przechodzącej przez punkt (4,5,6) oraz prostopadłej do: a) osi 0x, b) osi 0y, c) osi 0z.

6. Napisać równanie płaszczyzny przechodzącej przez punkt (3,-1,2) i zawierającej: a) oś 0x, b) oś 0y.

7. Napisać równanie płaszczyzny przechodzącej przez punkty (1,2,3) i (-1,3,4) oraz równoległej do: a) osi 0x, b) osi 0z.

8. Znaleźć punkty przecięcia płaszczyzny π z osiami układu współrzędnych 0xyz, gdy:
   a) π: $2x+y+3z-6=0$, b) π: $2x-y-3z=0$,

   c) π: $2x+y-6=0$, d) π: $2x+3z-6=0$.

9. Napisać równanie płaszczyzny przechodzącej przez punkty P_1, P_2, P_3:
   a) $P_1=(5,2,1)$, $P_2=(0,3,4)$, $P_3=(5,6,7)$,
   b) $P_1=(0,0,-12)$, $P_2=(2,-2,-5)$, $P_3=(4,0,-6)$,
   c) $P_1=(4,4,3)$, $P_2=(0,6,0)$, $P_3=(8,1,6)$.

10. Znaleźć wartości parametru k, dla których płaszczyzny π_1 i π_2 są równoległe, gdy:
    a) π_1: $2x+ky+z+6=0$, π_2: $kx+2y+(k-1)z+3=0$,
    b) π_1: $x+(k+1)y+6z+1=0$, π_2: $(k+1)x+4ky+(11+k^2)z=0$.

11. Dla jakich wartości parametru k płaszczyzny π_1 i π_2 z zadania 10 są prostopadłe?

12. Sprawdzić, że płaszczyzny π_1 i π_2 są równoległe, a następnie obliczyć odległość d tych płaszczyzn, gdy:
    a) π_1: $6x-3y+6z+5=0$, π_2: $4x-2y+4z-3=0$,
    b) π_1: $6x-8z-1=0$, π_2: $9x-12z+48=0$,
    c) π_1: $2x-4y-6z-2=0$, π_2: $3x-6y-9z-3=0$.

13. Napisać równanie płaszczyzny π zawierającej krawędź przecięcia płaszczyzn π_1 i π_2 i przechodzącej przez punkt P, gdy:
    a) π_1: $2x-3y-z-8=0$, π_2: $x-y-3z-6=0$, $P=(1,0,2)$,
    b) π_1: $x-3z-6=0$, π_2: $x+y-z-6=0$, $P=(1,2,3)$,
    c) π_1: $x+y-2z=0$, π_2: $y+2z-8=0$, $P=(0,2,1)$,
    d) π_1: $2x+2y+z-2=0$, π_2: $3x-y-z-2=0$, $P=(1,-1,2)$.

14. Napisać równanie płaszczyzny π zawierającej krawędź przecięcia płaszczyzn π_1 i π_2 i prostopadłej do płaszczyzny π_3, gdy:
    a) π_1: $x-y-3z-6=0$, π_2: $2x-3y-z-8=0$, π_3: $x+y-6z-12=0$,
    b) π_1: $2x-y-3=0$, π_2: $3y+z-8=0$, π_3: $x+y-6z-12=0$,
    c) π_1: $x+y-z-3=0$, π_2: $2x-y-3z-8=0$, π_3: $2x-y+z-6=0$,
    d) π_1: $x+y-z-3=0$, π_2: $2x-y-3z-8=0$, π_3: $4x-y+3z=0$.

15. Wyznaczyć zbiór punktów $(x,y,z)\in R^3$ określony równaniem:
a) $(2x-y-z+3)(x+3y-2)=0$,
b) $(2x-y-z+3)^2-(x+3y-2)^2=0$,
c) $(2x-y-z+3)^2+(x+3y-2)^2=0$.

Odpowiedzi.

2. a) $x-y-2z-3=0$, b) $x+3z=0$, c) $z=0$, d) $y-3=0$.
3. a) $x-y-z-3=0$, b) $3x-5y+2z=0$, c) $2x+3y-8=0$.
4. a) $5x+y+2z-15=0$, b) $3x-3y-z=0$, c) $y-3=0$, d) $x-1=0$.
5. a) $x-4=0$, b) $y-5=0$, c) $z-6=0$.
6. a) $2y+z=0$, b) $2x-3z=0$.
7. a) $y-z+1=0$, b) $x+2y-5=0$.
8. a) (3,0,0), (0,6,0), (0,0,2), b) π przecina osie jedynie w (0,0,0),
c) (3,0,0), (0,6,0), nie przecina osi 0z, d) (3,0,0), osi 0y nie przecina, (0,0,2).
9. a) $3x-15y+10z+5=0$, b) $3x-4y-2z-24=0$, c) $3x-4z=0$.
10. a) $k=2$, b) $k=1$.
11. a) $k=1/5$, b) nie istnieje k.
12. a) $d=19/18$; znaleźć dowolny punkt płaszczyzny π_1 np. $(0,0,-5/6)$ i obliczyć odległość P od π, b) $d=33/10$, c) $d=0$, płaszczyzny pokrywają się.
13. a) $14x-25y+13z-40=0$, b) $4x+7y+2z-24=0$,
c) $x+y-2z=0$, czyli szukaną płaszczyzną jest π_1,
d) takich płaszczyzn jest nieskończenie wiele; każda płaszczyzna
$\pi:\lambda_1(2x+2y+z-2)+\lambda_2(3x-y-z-2)=0$, gdzie $\lambda_1^2+\lambda_2^2>0$, spełnia podane warunki.
14. a) $31x-49y-3z-114=0$, b) $6x+z-17=0$,
c) $x+y-z-3=0$, czyli płaszczyzna π_1, d) każda płaszczyzna o równaniu
$\lambda_1(x+y-z-3)+\lambda_2(2x-y-3z-8)=0$, gdzie $\lambda_1^2+\lambda_2^2>0$, jest płaszczyzną prostopadłą do płaszczyzny π_3.
15. a) Suma $\pi_1\cup\pi_2$ płaszczyzn $\pi_1: 2x-y-z+3=0$, $\pi_2: x+3y-2=0$,
b) suma $\pi_1\cup\pi_2$ płaszczyzn $\pi_1: x-4y-z+5=0$, $\pi_2: 3x+2y-z+1=0$; gdyż
$(2x-y-z+3)^2=(x+3y-2)^2 \Leftrightarrow 2x-y-z+3=x+3y-2$ lub
$2x-y-z+3=-(x+3y-2)$,
c) iloczyn $\pi_1\cap\pi_2$ płaszczyzn $\pi_1: 2x-y-z+3=0$, $\pi_2: x+3y-2=0$, czyli krawędź przecięcia tych płaszczyzn.

3. PROSTA W PRZESTRZENI.

RÓWNANIA PARAMETRYCZNE I KIERUNKOWE PROSTEJ. Znajdziemy równania prostej mając dany punkt leżący na tej prostej i wektor do niej równoległy.

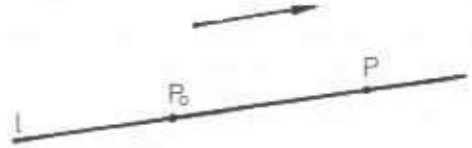
TWIERDZENIE 3.1. Załóżmy, że niezerowy wektor $\bar{r}=[a,b,c]$ $(\text{tzn. } a^2+b^2+c^2>0)$ jest równoległy do prostej l i punkt $P_0(x_0,y_0,z_0)$ należy do tej prostej (rys 3.1).

Wówczas punkt $P(x,y,z)$ należy do prostej l wtedy i tylko wtedy, gdy liczby x, y, z spełniają układ równań

$$(3.1)\qquad \begin{cases} x=x_0+at, \\ y=y_0+bt, \\ z=z_0+ct, \end{cases} \quad t\in R.$$

Równania (3.1) nazywamy **równaniami parametrycznymi** prostej l, **t-parametrem**, zaś a, b, c - **współczynnikami kierunkowymi** tej prostej; równania (3.1) będziemy również zapisywać w postaci

$$x=x_0+at,\quad y=y_0+bt,\quad z=z_0+ct,\quad t\in R.$$



Rys 3.1.

D o w ó d. Niech $P(x,y,z)$ będzie dowolnym różnym od P_0 punktem w układzie współrzędnych 0xyz. Wówczas

$$P(x,y,z)\in l \Leftrightarrow \overrightarrow{P_0P}\parallel\bar{r} \Leftrightarrow \overrightarrow{P_0P}=t\bar{r},\ t\in R \Leftrightarrow$$
$$\Leftrightarrow [x-x_0,y-y_0,z-z_0]=t[a,b,c],\ t\in R \Leftrightarrow$$
$$\Leftrightarrow (x-x_0=at,\ y-y_0=bt,\ z-z_0=ct,\ t\in R).$$

Otrzymany układ trzech równań równoważny jest układowi (3.1). Łatwo sprawdzić, że współrzędne x_0, y_0, z_0 punktu P_0 również spełniają ten układ.

Jeśli $a \neq 0$, $b \neq 0$, $c \neq 0$ (czyli $abc \neq 0$), to równania parametryczne można zapisać w postaci

$$\frac{x-x_0}{a}=t, \quad \frac{y-y_0}{b}=t, \quad \frac{z-z_0}{c}=t, \qquad t \in R,$$

skąd otrzymujemy

$$(3.2) \qquad l: \ \frac{x-x_0}{a}=\frac{y-y_0}{b}=\frac{z-z_0}{c}.$$

Równania (3.2) nazywamy **równaniami kierunkowymi** prostej.

Dla przykładu: 1) Prosta l przechodząca przez punkt $P_0(1,2,3)$ i równoległa do wektora $\vec{r}=[4,3,-1]$ ma równania parametryczne

$$l: \ x=1+4t, \ y=2+3t, \ z=3-t, \ \ t \in R.$$

Ponieważ wszystkie współrzędne wektora $\vec{r}$ są różne od zera, więc prosta ta ma również równania kierunkowe

$$l: \ \frac{x-1}{4}=\frac{y-2}{3}=\frac{z-3}{-1}.$$

2) Prosta l przechodząca przez punkt $P_0(-1,2,3)$ i równoległa do osi 0x ma równania parametryczne

$$l: \ x=-1+t, \ y=2, \ z=3, \ \ t \in R,$$

ponieważ wektorem równoległym do tej prostej jest wersor osi 0x: $\vec{r}=\vec{i}=[1,0,0]$. Dla tej prostej nie istnieją równania kierunkowe, gdyż nie wszystkie współrzędne wektora $\vec{r}$ są różne od zera.

3) Oś 0x ma równania parametryczne

$$x=t, \ y=0, \ z=0, \ \ t \in R,$$

gdyż jest to prosta przechodząca przez punkt O(0,0,0) i równoległa do wektora $\vec{r}=\vec{i}=[1,0,0]$.

Analogicznie: równania

$$x=0, \ y=t, \ z=0, \ \ t \in R,$$

są równaniami parametrycznymi osi 0y, a równania

$$x=0, \ y=0, \ z=t, \ \ t \in R,$$

są równaniami parametrycznymi osi 0z.

PRZYKŁAD 3.1 (równania prostej przechodzącej przez dwa punkty). Napiszemy równania parametryczne prostej l przechodzącej przez punkty $P(-1,3,0)$ i $Q(4,5,7)$.

Punkt $P(-1,3,0)$ należy do prostej l, zatem

$$l: \ x=-1+at, \ y=3+bt, \ z=ct, \ \ t \in R,$$

gdzie wektor $\vec{r}=[a,b,c]$ jest dowolnym wektorem równoległym do prostej l. Zatem możemy przyjąć $\vec{r}=\overrightarrow{PQ}$, czyli

$$[a,b,c]=[5,2,7].$$

Prosta l ma więc równania parametryczne

$$l: \ x=-1+5t, \ y=3+2t, \ z=7t, \ \ t \in R. \qquad \blacksquare$$

U w a g a . Jeśli w tym przykładzie przyjmiemy

$$(x_0,y_0,z_0)=(4,5,7), \ \ \vec{r}=\overrightarrow{QP},$$

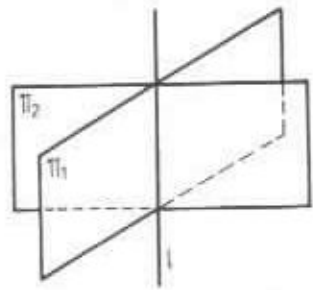
to otrzymamy inne równania parametryczne tej samej prostej:

$$l: \ x=4-5t, \ y=5-2t, \ z=7-7t, \ \ t \in R.$$

Widzimy więc, że linia może być opisana różnymi równaniami parametrycznymi.

RÓWNANIA KRAWĘDZIOWE PROSTEJ. Niech l oznacza prostą, która jest krawędzią przecięcia dwu nierównoległych płaszczyzn (rys 3.2)

$$\pi_1: \ A_1x+B_1y+C_1z+D_1=0, \quad \pi_2: \ A_2x+B_2y+C_2z+D_2=0.$$



Rys 3.2.

Punkt $P(x,y,z) \in l$ wtedy i tylko wtedy, gdy $P \in \pi_1$ i $P \in \pi_2$, czyli wtedy i tylko wtedy, gdy liczby x, y, z spełniają układ równań liniowych

$$(3.3)\qquad \begin{cases} A_1x+B_1y+C_1z+D_1=0, \\ A_2x+B_2y+C_2z+D_2=0. \end{cases}$$

Równania (3.3) nazywamy **równaniami krawędziowymi** prostej l.

U w a g a . Wektor $\vec{r}=[a,b,c]$ równoległy do prostej l danej równaniami krawędziowymi (3.3) jest równoległy do każdej z płaszczyzn π_1 i π_2, więc

$$\vec{r}\perp[A_1,B_1,C_1], \qquad \vec{r}\perp[A_2,B_2,C_2],$$

a zatem możemy przyjąć

$$\vec{r}=[a,b,c]=[A_1,B_1,C_1]\times[A_2,B_2,C_2].$$

Stąd wynika, że współczynniki kierunkowe a, b, c prostej danej równaniami krawędziowymi (3.3) wyrażają się wzorami

$$(3.4)\qquad a=\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}, \quad b=\begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix}=-\begin{vmatrix} A_1 & C_1 \\ A_2 & C_2 \end{vmatrix}, \quad c=\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}.$$

P R Z Y K Ł A D 3.2. Napiszemy równania parametryczne prostej l danej równaniami krawędziowymi

$$(1)\qquad l:\begin{cases} 2x-3y-3z=0, \\ x-6y-6z+9=0. \end{cases}$$

I s p o s ó b . Należy znaleźć dowolny punkt $P_0(x_0,y_0,z_0)\in l$ i wektor $\vec{r}=[a,b,c]\|l$. Aby wyznaczyć punkt P_0 prostej l trzeba znaleźć dowolne trzy liczby x, y, z spełniające układ równań (1). Przyjmując $x=0$ w równaniach (1) otrzymujemy układ sprzeczny

$$\begin{cases} y+z=0, \\ 2y+2z=3, \end{cases}$$

co oznacza, że do prostej l nie należy punkt, którego współrzędna $x=0$. Przyjmijmy więc $y=0$. Wówczas z równań krawędziowych (1) prostej l otrzymamy układ równań

$$\begin{cases} 2x-3z=0, \\ x-6z+9=0, \end{cases}$$

którego rozwiązaniem jest para liczb $x=3$, $z=2$. Zatem punkt $P_0(3,0,2)$ należy do prostej l.

Wektor [a,b,c] równoległy do prostej l jest równoległy do płaszczyzn $\pi_1: 2x-3y-3z=0$, $\pi_2: x-6y-6z+9=0$; spełnia więc warunki

$$[a,b,c]\perp[2,-3,-3], \qquad [a,b,c]\perp[1,-6,-6].$$

Możemy zatem przyjąć

$$[a,b,c]=[2,-3,-3]\times[1,-6,-6]=[0,9,-9].$$

Prosta l ma równania parametryczne

$$l:\ x=3,\ y=9t,\ z=2-9t, \quad t\in R.$$

Współczynniki kierunkowe a, b, c prostej l można oczywiście obliczyć korzystając bezpośrednio ze wzorów (3.4):

$$a=\begin{vmatrix} -3 & -3 \\ -6 & -6 \end{vmatrix}=0,\ b=\begin{vmatrix} -3 & 2 \\ -6 & 1 \end{vmatrix}=9,\ c=\begin{vmatrix} 2 & -3 \\ 1 & -6 \end{vmatrix}=-9.$$

II s p o s ó b . Przyjmijmy np. $y=t$, $t\in R$. Z równań krawędziowych prostej l otrzymujemy układ dwu równań liniowych

$$\begin{cases} 2x-3t-3z=0, \\ x-6t-6z+9=0, \end{cases}$$

Z tego układu otrzymujemy $x=3$, $z=2-t$, a zatem

$$l:\ x=3,\ y=t,\ z=2-t, \quad t\in R.$$

Otrzymaliśmy tutaj inne niż poprzednio równania parametryczne tej samej prostej l (zob. również uwagę po przykładzie 3.1). ■

WZAJEMNE POŁOŻENIE DWU PROSTYCH W PRZESTRZENI. Położenie dwu prostych w przestrzeni będziemy rozpatrywać ze względu na kąt między tymi prostymi.

Przypomnijmy, że miara kąta między prostymi jest liczbą z przedziału $<0,\pi/2>$.

TWIERDZENIE 3.2 Załóżmy, że proste l_1 i l_2 mają równania parametryczne

$$l_1:\begin{cases} x=x_0'+a_1t, \\ y=y_0'+b_1t, \\ z=z_0'+c_1t, \end{cases} \quad t\in R, \qquad l_2:\begin{cases} x=x_0''+a_2s, \\ y=y_0''+b_2s, \\ z=z_0''+c_2s, \end{cases} \quad s\in R.$$

Wówczas

$$(1)\qquad \cos\measuredangle(l_1,l_2)=\frac{|a_1a_2+b_1b_2+c_1c_2|}{\sqrt{a_1^2+b_1^2+c_1^2}\sqrt{a_2^2+b_2^2+c_2^2}},$$

(2) $$\sin\measuredangle(l_1,l_2)=\frac{\|[a_1,b_1,c_1]\times[a_2,b_2,c_2]\|}{\sqrt{a_1^2+b_1^2+c_1^2}\,\sqrt{a_2^2+b_2^2+c_2^2}},$$

(3) $$l_1\perp l_2 \Leftrightarrow a_1a_2+b_1b_2+c_1c_2=0,$$

(4) $$l_1\parallel l_2 \Leftrightarrow \frac{a_1}{a_2}=\frac{b_1}{b_2}=\frac{c_1}{c_2}.$$

PRZYKŁAD 3.3. Napiszemy równania parametryczne prostej l przechodzącej przez punkt (2,0,1) i równoległej do prostej

$$l_1:\ \frac{x-1}{3}=2y-4=\frac{z}{6}.$$

Ponieważ $(2,0,1)\in l$, więc

$$l:\begin{cases}x=2+at,\\ y=bt,\\ z=1+ct,\end{cases}\quad t\in R.$$

Z warunku równoległości prostych l i l_1 wynika, że jako wektor [a,b,c] równoległy do prostej l możemy przyjąć wektor równoległy do prostej l_1. Ponieważ

$$\frac{x-1}{3}=2y-4=\frac{z}{6} \Leftrightarrow \frac{x-1}{3}=\frac{y-2}{1/2}=\frac{z}{6},$$

więc

$$[a,b,c]=[3,1/2,6],$$

a zatem

$$l:\begin{cases}x=2+3t,\\ y=t/2,\\ z=1+6t,\end{cases}\quad t\in R.$$ ■

PRZYKŁAD 3.4. Napiszemy równania parametryczne prostej l przechodzącej przez punkt (3,2,-1) i prostopadłej do prostych

$$l_1:\begin{cases}x=t,\\ y=4,\\ z=1+2t,\end{cases}\quad t\in R,\qquad l_2:\ 2x=y=\frac{z-6}{3}.$$

Ponieważ $(3,2,-1)\in l$, więc

$$l:\begin{cases}x=3+at,\\ y=2+bt,\\ z=-1+ct,\end{cases}\quad t\in R.$$

Wektor [a,b,c] równoległy do prostej l jest prostopadły do prostych l_1 i l_2, a zatem

$$[a,b,c]\perp[1,0,2],\qquad [a,b,c]\perp[1/2,1,3].$$

Możemy więc przyjąć

$$[a,b,c]=[1,0,2]\times[1/2,1,3]=[-2,-2,1].$$

Zatem prosta l ma równania parametryczne

$$l:\begin{cases}x=3-2t,\\ y=2-2t,\\ z=-1+t,\end{cases}\quad t\in R.$$ ■

W następnych przykładach będziemy badać wzajemne położenie dwu prostych. Przypomnijmy,że dwie proste w przestrzeni mogą przecinać się, być równoległe (w szczególności pokrywać się) lub być skośne.

PRZYKŁAD 3.5. Zbadamy wzajemne położenie prostych

$$l_1:\begin{cases}x=2+3t,\\ y=4,\\ z=3+6t,\end{cases}\quad t\in R,\qquad l_2:\begin{cases}x=3+s,\\ y=4,\\ z=2s,\end{cases}\quad s\in R.$$

Wektory [3,0,6], [1,0,2] są równoległe, więc te proste są również równoległe. Zatem proste te mają nieskończenie wiele punktów wspólnych (pokrywają się) lub nie mają ani jednego punktu wspólnego. Punkt (2,4,3) należy do prostej l_1. Sprawdzamy, czy punkt ten należy również do prostej l_2 tzn. czy istnieje taka wartość parametru s, by spełnione były równania

$$\begin{cases}2=3+s,\\ 4=4,\\ 3=2s.\end{cases}$$

Stąd otrzymujemy sprzeczną koniunkcję: $s=-1 \wedge s=\frac{3}{2}$. Zatem proste l_1 i l_2 są równoległe i nie mają punktów wspólnych. ■

PRZYKŁAD 3.6. Zbadamy wzajemne położenie prostych

$$l_1:\begin{cases}x=2+3t,\\ y=4+9t,\\ z=3+6t,\end{cases}\quad t\in R,\qquad l_2:\begin{cases}x=3+2s,\\ y=4+6s,\\ z=3+2s,\end{cases}\quad s\in R.$$

Ponieważ wektory [3,9,6], [2,6,2] nie są równoległe, więc te proste też nie są równoległe. Mogą być skośne lub przecinać się. Jeżeli proste l_1 i l_2 przecinają się, to istnieją takie wartości parametrów t i s, że

$$2+3t=3+2s,\quad 4+9t=4+6s,\quad 3+6t=3+2s.$$

Ponieważ otrzymany układ równań jest sprzeczny, więc proste l_1 i l_2 nie przecinają się. Proste l_1 i l_2 są skośne. ■

ZADANIA DO ROZWIĄZANIA.

1. Jakie elementy wyznaczają prostą w przestrzeni R^3?
2. Czy punkt i wektor równoległy do prostej wyznaczają prostą w przestrzeni: a) R^2, b) R^3?
3. Czy punkt i wektor prostopadły do prostej wyznaczają prostą w przestrzeni: a) R^2, b) R^3?
4. Napisać równania parametryczne prostej przechodzącej przez punkt (3,4,-2) i równoległej do osi: a) 0x, b) 0y, c) 0z.
5. Napisać równania parametryczne prostej l przechodzącej przez punkt (3,4,-2) i równoległej do prostej l_1, gdy:

   a) $l_1:\begin{cases}x=t,\\y=3,\\z=2-t,\end{cases}\ t\in R$, b) $l_1: x=4y=\frac{z-6}{3}$, c) $l_1:\begin{cases}2x-y-z-6=0,\\x+y+3z-5=0.\end{cases}$
6. Napisać równania parametryczne prostej l przechodzącej przez punkty P(3,4,-2) i Q(5,6,2) i sprawdzić, czy punkt R(1,-2,3) należy do tej prostej.
7. Napisać równania parametryczne prostej przechodzącej przez punkt (3,4,5) i przecinającej oś 0y w punkcie o współrzędnej $y=5$.
8. Napisać równania parametryczne prostej przechodzącej przez punkt (-1,2,3) i prostopadłej do prostych l_1 i l_2, jeśli:

   a) $l_1:\begin{cases}x=2+3t,\\y=t,\\z=3,\end{cases}\ t\in R,\quad l_2:\begin{cases}2x-y+2z-6=0,\\x+y+3z-4=0,\end{cases}$

   b) $l_1:\ \frac{x-1}{3}=\frac{2-y}{3}=2z,\quad l_2:\begin{cases}x+3y-6=0,\\2x-y-z-8=0.\end{cases}$
9. Napisać równania parametryczne, kierunkowe i krawędziowe prostej przechodzącej przez punkty P(1,2,0) i Q(−1,3,−4).
10. Sprawdzić, że proste l_1 i l_2 są równoległe, jeśli

    a) $l_1:\ x-1=2y=\frac{z-3}{2}$, $\quad l_2:\begin{cases}4x+12y-5z=0,\\4x+4y-3z+1=0,\end{cases}$

    b) $l_1:\ x=3y=z-1$, $\quad l_2:\ \frac{x+3}{3}=y+1=\frac{z+2}{3}$.
11. Ile punktów wspólnych maja proste l_1 i l_2 z zadania 10?
12. Znaleźć (jeśli istnieją) punkty wspólne prostych l_1, l_2, gdy

    a) $l_1:\ x=3y=z-1$, $\quad l_2:\ \frac{x-2}{3}=-y=\frac{z-1}{2}$,

    b) $l_1:\ x=3y=z-1$, $\quad l_2:\ \frac{x-2}{3}=-y=\frac{z-4}{6}$.

Odpowiedzi.

2. a) Tak, b) tak, 3. a) Tak, b) nie.
4. a) $\begin{cases}x=3+t,\\y=4,\\z=-2,\end{cases}$ b) $\begin{cases}x=3,\\y=4+t,\\z=-2,\end{cases}$ c) $\begin{cases}x=3,\\y=4,\\z=-2+t.\end{cases}$
5. a) $\begin{cases}x=3+t,\\y=4,\\z=-2-t,\end{cases}$ b) $\begin{cases}x=3+t,\\y=4+t/4,\\z=-2+3t,\end{cases}$ c) $\begin{cases}x=3-2t,\\y=4-7t,\\z=-2+3t.\end{cases}$
6. $\begin{cases}x=3+2t,\\y=4+2t,\\z=-2+4t,\end{cases}$ nie. 7. $\begin{cases}x=3t,\\y=5-t,\\z=5t.\end{cases}$
8. a) $\begin{cases}x=-1-3t,\\y=2+9t,\\z=3+7t,\end{cases}$ b) $\begin{cases}x=-1+4t,\\y=2+39t,\\z=3-12t.\end{cases}$
9. $\begin{cases}x=1+2t,\\y=2-t,\\z=4t,\end{cases}$ $\quad \frac{x-1}{2}=\frac{y-2}{-1}=\frac{z}{4}$, $\quad \begin{cases}x+2y-5=0\\4y+z-8=0.\end{cases}$
11. a) Nie mają punktów wspólnych, b) mają nieskończenie wiele punktów wspólnych, gdyż proste pokrywają się.
12. a) Nie maja punktów wspólnych, są skośne, b) przecinają się w punkcie (1,1/3,2).

4. PROSTA I PŁASZCZYZNA.

WZAJEMNE POŁOŻENIE PROSTEJ I PŁASZCZYZNY.

Położenie prostej i płaszczyzny względem siebie będziemy rozpatrywać ze względu na kąt między prostą i płaszczyzną.

Przypomnijmy, że miara kąta między prostą i płaszczyzną jest liczbą z przedziału $<0,\frac{\pi}{2}>$.

TWIERDZENIE 4.1. Załóżmy, że prosta l i płaszczyzna π mają odpowiednio równania

$$l:\begin{cases} x = x_0 + at, \\ y = y_0 + bt, \\ z = z_0 + ct, \end{cases} \quad t \in R, \qquad \pi:\ Ax + By + Cz + D = 0.$$

Wówczas

(1) $l \parallel \pi \Leftrightarrow aA + bB + cC = 0,$

(2) $l \perp \pi \Leftrightarrow \frac{a}{A} = \frac{b}{B} = \frac{c}{C},$

(3) $\cos\sphericalangle(l,\pi) = \frac{|[a,b,c]\times[A,B,C]|}{\sqrt{a^2+b^2+c^2}\,\sqrt{A^2+B^2+C^2}},$

(4) $\sin\sphericalangle(l,\pi) = \frac{|aA+bB+cC|}{\sqrt{a^2+b^2+c^2}\,\sqrt{A^2+B^2+C^2}}.$

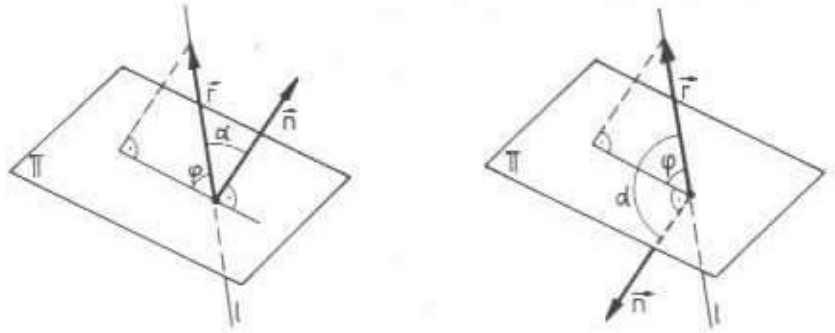
Dowód (1). Wektor [a,b,c] jest równoległy do prostej l, wektor [A,B,C] jest prostopadły do płaszczyzny π. Zatem prosta l i płaszczyzna π są równoległe wtedy i tylko wtedy, gdy wektory [a,b,c], [A,B,C] są prostopadłe, czyli

$$l \parallel \pi \Leftrightarrow [a,b,c] \perp [A,B,C] \Leftrightarrow aA + bB + cC = 0.$$

Dowód (4). Niech α oznacza kąt wektora $\vec{r} = [a,b,c]$ z wektorem $\vec{n} = [A,B,C]$ zaś φ - kąt między prostą l i płaszczyzną π (rys 4.1).

Można zauważyć (rys. 4.1), że

$$\varphi = \begin{cases} \frac{\pi}{2} - \alpha, & \text{gdy } \alpha \in <0,\pi/2>, \\ \alpha - \frac{\pi}{2}, & \text{gdy } \alpha \in (\pi/2,\pi>, \end{cases}$$



Rys 4.1.

wobec tego

$$\sin\varphi = \begin{cases} \cos\alpha, & \text{gdy } \alpha \in <0,\pi/2>, \\ -\cos\alpha, & \text{gdy } \alpha \in (\pi/2,\pi>. \end{cases}$$

Zatem $\sin\varphi = |\cos\alpha|$. Uwzględniając wzór na cosinus kąta α między wektorami $\vec{r} = [a,b,c]$ i $\vec{n} = [A,B,C]$ otrzymujemy tezę (4).

PRZYKŁAD 4.1. Napiszemy równanie płaszczyzny π, która przechodzi przez punkt P(-1,0,3) i jest prostopadła do prostej l: $x+1 = -y = z/3$.

Ponieważ $P(-1,0,3) \in \pi$, więc

$$\pi:\ A(x+1) + By + C(z-3) = 0.$$

Z warunków zadania wynika, że $\pi \perp l$. Ponieważ

$$\pi \perp l \Leftrightarrow [A,B,C] \parallel [1,-1,3],$$

więc możemy przyjąć [A,B,C] = [1,-1,3]. Zatem

$$\pi:\ (x+1) - y + 3(z-3) = 0,$$

czyli $\pi:\ x - y + 3z - 8 = 0$. ■

PRZYKŁAD 4.2. Napiszemy równanie płaszczyzny π przechodzącej przez punkt P(-1,0,3), równoległej do prostej l: $x = 2 - y = 2z$ i prostopadłej do płaszczyzny π_1: $2x + 4y - z - 8 = 0$.

I sposób. Ponieważ $P(-1,0,3) \in \pi$, więc

$$\pi:\ A(x+1) + By + C(z-3) = 0.$$

Z warunków zadania mamy: $\pi \| l$ i $\pi \perp \pi_1$. Ponieważ

$$\pi \| l \Leftrightarrow [A,B,C] \perp [1,-1,1/2],$$
$$\pi \perp \pi_1 \Leftrightarrow [A,B,C] \perp [2,4,-1],$$

więc możemy przyjąć

$$[A,B,C] = [1,-1,1/2] \times [2,4,-1] = [-1,2,6].$$

Zatem

$$\pi: -(x+1)+2y+6(z-3)=0,$$
$$\pi: x-2y-6z+19=0.$$

II sposób.

$$\left.\begin{array}{rl} P(-1,0,3) \in \pi \Leftrightarrow & \pi: A(x+1)+By+C(z-3)=0, \\ \pi \| l \Leftrightarrow & A\cdot 1+B(-1)+C\cdot 1/2=0, \\ \pi \perp \pi_1 \Leftrightarrow & A\cdot 2+B\cdot 4+C\cdot(-1)=0. \end{array}\right\}$$

Jest to układ równań jednorodnych z niewiadomymi A, B, C, przy czym wiadomo, że układ ten ma rozwiązania niezerowe, gdyż [A,B,C] jest wektorem niezerowym. Zatem z twierdzenia Cramera wynika, że wyznacznik główny tego układu równań jest równy zeru:

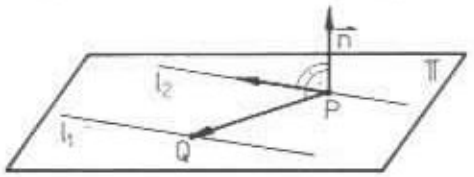
$$\begin{vmatrix} x+1 & y & z-3 \\ 1 & -1 & 1/2 \\ 2 & 4 & -1 \end{vmatrix} = 0.$$

Stąd otrzymujemy równanie płaszczyzny π

$$\pi: x-2y-6z+19=0. \quad \blacksquare$$

PRZYKŁAD 4.3. Wyznaczymy równanie płaszczyzny π zawierającej proste równoległe (rys 4.2)

$$l_1: \frac{x+3}{2} = 2-y = 2z, \quad l_2: \frac{x}{4} = \frac{3-y}{2} = z+1.$$



Rys 4.2.

Ponieważ $P(0,3,-1) \in l_2$, więc $P \in \pi$, a zatem

$$\pi: Ax+B(y-3)+C(z+1)=0.$$

Niech Q będzie dowolnie ustalonym punktem należącym do prostej l_1 np. $Q(-3,2,0)$. Wówczas wektor $\vec{n} = [A,B,C]$ spełnia warunki: $\vec{n} \perp \overrightarrow{PQ}$ i $\vec{n} \perp l_2$. Ponieważ

$$\vec{n} \perp \overrightarrow{PQ} \Leftrightarrow \vec{n} \perp [-3,-1,1],$$
$$\vec{n} \perp l_2 \Leftrightarrow \vec{n} \perp [4,-2,1],$$

więc możemy przyjąć

$$\vec{n} = [A,B,C] = [-3,-1,1] \times [4,-2,1] = [1,7,10].$$

W konsekwencji otrzymujemy

$$\pi: x+7(y-3)+10(z+1)=0,$$

czyli $\pi: x+7y+10z-11=0$. $\blacksquare$

PRZYKŁAD 4.4. Znajdziemy punkt przecięcia prostej l z płaszczyzną π, gdy

$$l: \frac{x-1}{2} = y = 3z, \qquad \pi: x+2y+z+12=0.$$

Punkt wspólny prostej i płaszczyzny znajdujemy rozwiązując układ równań

$$\begin{cases} \dfrac{x-1}{2} = y = 3z, \\ x+2y+z+12=0 \end{cases}$$

lub równoważny mu układ równań

$$\begin{cases} x = 1+2t, \\ y = t, \\ z = t/3, \\ x+2y+z+12=0. \end{cases}$$

Z tego ostatniego układu otrzymujemy

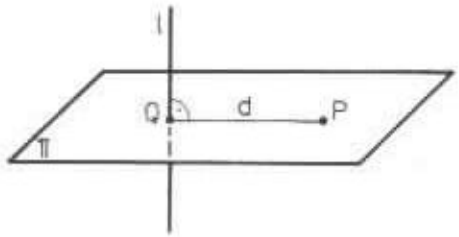
$$1+2t+2t+t/3+12=0,$$

czyli $t=-3$, a stąd współrzędne punktu przecięcia prostej z płaszczyzną:

$$\begin{cases} x = 1+2\cdot(-3) = -5, \\ y = -3, \\ z = (-3)\cdot 1/3 = -1. \end{cases}$$

Punkt P(-5,-3,-1) jest więc punktem przecięcia prostej l z płaszczyzną π. $\blacksquare$

PRZYKŁAD 4.5. Obliczymy odległość d punktu P(2,3,5) od prostej l: $x=3-y=2z$ (rys 4.3).



Rys 4.3.

Piszemy najpierw równanie płaszczyzny π przechodzącej przez punkt P i prostopadłej do prostej l:

$$P(2,3,5)\in\pi \Leftrightarrow \pi:\ A(x-2)+B(y-3)+C(z-5)=0,$$

$$\pi\perp l \Leftrightarrow \frac{A}{1}=\frac{B}{-1}=\frac{C}{1/2}.$$

Przyjmując $A=1$, obliczamy $B=-1$, $C=1/2$ i otrzymujemy

$$\pi:\ 2x-2y+z-3=0.$$

Następnie znajdujemy punkt Q, który jest punktem wspólnym płaszczyzny π i prostej l. Rozwiązując układ równań

$$\begin{cases} x=3-y=2z, \\ 2x-2y+z-3=0 \end{cases}$$

otrzymujemy, że Q=(2,1,1).

Obliczamy odległość punktów P i Q. Jest to szukana odległość punktu P od prostej l:

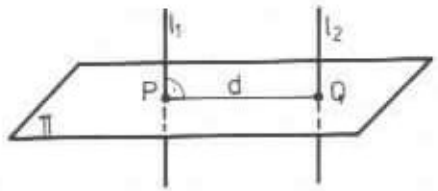
$$d=d(P,l)=|PQ|=\sqrt{(2-2)^2+(1-3)^2+(1-5)^2}=2\sqrt{5}. \quad \blacksquare$$

PRZYKŁAD 4.6. Obliczymy odległość d prostych równoległych (rys. 4.4):

$$l_1:\ x+1=2y=z, \qquad l_2:\ \frac{x}{2}=y-3=\frac{z+2}{2}.$$

Piszemy równanie płaszczyzny π przechodzącej przez punkt $P(-1,0,0)\in l_1$ i prostopadłej do prostych l_1 i l_2:

$$\pi:\ 2x+y+2z+2=0.$$



Rys 4.4.

Współrzędne punktu Q, punktu wspólnego płaszczyzny π i prostej l_2, znajdujemy rozwiązując układ równań

$$\begin{cases} \frac{x}{2}=y-3=\frac{z+2}{2}, \\ 2x+y+2z+2=0. \end{cases}$$

Zatem Q=(−2/9,26/9,−20/9).

Obliczamy odległość punktów P i Q. Jest to szukana odległość prostych l_1 i l_2:

$$d=d(l_1,l_2)=|PQ|=\sqrt{(-\frac{2}{9}+1)^2+(\frac{26}{9})^2+(-\frac{20}{9})^2}=\frac{5\sqrt{5}}{3}. \quad \blacksquare$$

PRZYKŁAD 4.7. Obliczymy odległość prostych

l_1: $x+1=2y=z$, l_2: $x/2=y-1=z/6$.

Wektory $\vec{r}_1=[1,1/2,1]$ i $\vec{r}_2=[2,1,6]$ nie są równoległe, zatem proste l_1 i l_2 nie są równoległe.

Piszemy równanie płaszczyzny π zawierającej prostą l_2 i równoległej do prostej l_1:

$$P_0(0,1,0)\in l_2 \Rightarrow P_0\in\pi \Leftrightarrow \pi:\ Ax+B(y-1)+Cz=0,$$

$$\pi\parallel l_2 \Leftrightarrow [A,B,C]\perp[2,1,6] \Leftrightarrow 2A+B+6C=0,$$

$$\pi\parallel l_1 \Leftrightarrow [A,B,C]\perp[1,1/2,1] \Leftrightarrow A+\frac{1}{2}B+C=0.$$

Stąd otrzymujemy

$$\pi:\ \begin{vmatrix} x & y-1 & z \\ 2 & 1 & 6 \\ 1 & 1/2 & 1 \end{vmatrix}=0,$$

czyli $\pi:\ x-2y+2=0$.

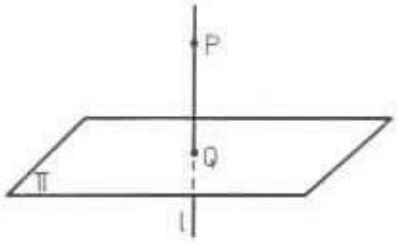
Wybieramy dowolny punkt prostej l_1, np. P(-1,0,0) i obliczamy jego odległość od płaszczyzny π. Jest to odległość prostych l_1 i l_2:

$$d = d(l_1,l_2)=d(P,\pi)=\frac{|-1-2\cdot 0+2|}{\sqrt{1^2+(-2)^2}}=\frac{1}{\sqrt{5}}.$$ ■

PRZYKŁAD 4.8. Znajdziemy punkt Q,który jest rzutem prostokątnym punktu $P(-3,1,-2)$ na płaszczyznę $\pi: 2x-y+3z-1=0$.

Najpierw piszemy równania prostej l przechodzącej przez punkt P i prostopadłej do płaszczyzny π (rys. 4.5):

$$l:\ x=-3+2t,\quad y=1-t,\quad z=-2+3t,\quad t\in R.$$



Rys 4.5

Rzutem prostokątnym punktu P na płaszczyznę π jest punkt przecięcia prostej l z płaszczyzną π. Rozwiązujemy więc układ równań

$$\begin{cases} x=-3+2t,\ y=1-t,\ z=-2+3t,\\ 2x-y+3z-1=0. \end{cases}$$

Stąd otrzymujemy, że $Q=(-1,0,1)$. ■

PRZYKŁAD 4.9. Napiszemy równania parametryczne prostej l_1, która jest rzutem prostokątnym prostej $l: \frac{x+1}{2}=1-y=\frac{z}{2}$ na płaszczyznę $\pi: x-2z+3=0$.

Wybieramy dwa dowolne punkty prostej l, np. $P_1(-1,1,0)$ i $P_2(1,0,2)$. Piszemy równanie płaszczyzny π_1 przechodzącej przez prostą l i prostopadłej do płaszczyzny π:

$$P_1(-1,1,0)\in\pi_1 \Leftrightarrow \pi_1: A(x+1)+B(y-1)+Cz=0,$$
$$P_2(1,0,2)\in\pi_1 \Leftrightarrow 2A-B+2C=0,$$
$$\pi_1\perp\pi \Leftrightarrow A-2C=0.$$

Stąd

$$\pi_1:\begin{vmatrix} x+1 & y-1 & z\\ 2 & -1 & 2\\ 1 & 0 & -2 \end{vmatrix}=0,$$

czyli

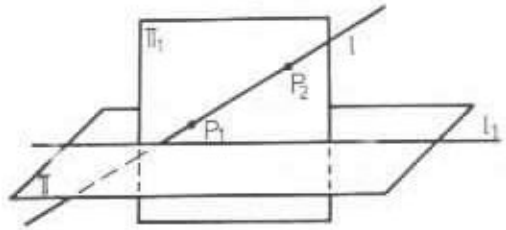
$$\pi_1: 2x+6y+z-4=0.$$

Rzutem prostokątnym prostej l na płaszczyznę π jest krawędź przecięcia płaszczyzn π i π_1 (rys. 4.6), czyli

$$l_1:\begin{cases} x-2z+3=0,\\ 2x+6y+z-4=0. \end{cases}$$

Przyjmując np. $z=t$ otrzymujemy stąd równania parametryczne prostej l_1

$$l_1:\ x=-3+2t,\ y=\frac{5}{3}-\frac{5}{6}t,\ z=t,\quad t\in R.$$ ■



Rys 4.6

PRZYKŁAD 4.10. Znajdziemy punkt P symetryczny do punktu $Q(0,4,6)$ względem płaszczyzny $\pi: 2x-y-3z+8=0$.

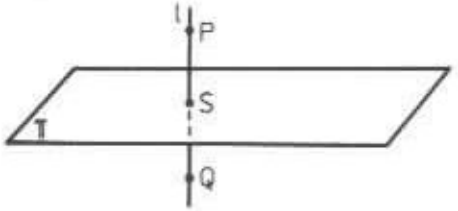
Piszemy najpierw równania prostej l przechodzącej przez Q i prostopadłej do płaszczyzny π (rys. 4.7)

$$l:\ \frac{x}{2}=\frac{y-4}{-1}=\frac{z-6}{-3}.$$

Następnie znajdujemy współrzędne punktu S, który jest punktem wspólnym prostej l i płaszczyzny π; rozwiązujemy więc układ równań

$$\begin{cases} \frac{x}{2}=4-y=\frac{z-6}{-3},\\ 2x-y-3z+8=0. \end{cases}$$

Stąd otrzymujemy, że $S=(2,3,3)$.



Rys 4.7.

Znajdujemy współrzędne x, y, z punktu P wiedząc, że punkt S jest środkiem odcinka $\overline{PQ}$. Ponieważ współrzędne punktu S są średnimi arytmetycznymi odpowiednich współrzędnych punktów P i Q, więc

$$2=\frac{x+0}{2},\quad 3=\frac{y+4}{2},\quad 3=\frac{z+6}{2},$$

a stąd $x=4$, $y=2$, $z=0$. Punktem symetrycznym do punktu Q względem płaszczyzny π jest więc punkt P(4,2,0). ■

ZADANIA DO ROZWIĄZANIA.

1. Napisać równanie płaszczyzny przechodzącej przez punkt P(1,0,2) i prostopadłej do prostej l, jeśli:

   a) $l\colon \frac{x}{3}=3-y=6z$, b) $l\colon \begin{cases}2x-y-z=0\\ x+3z-1=0\end{cases}$, c) $l\colon \begin{cases}x=t,\\ y=3,\\ z=2-3t.\end{cases}$

2. Napisać równania prostej l przechodzącej przez punkt P(1,0,2) i prostopadłej do płaszczyzny π, jeśli:

   a) $\pi\colon x-3y-z-6=0$, b) $\pi\colon x-3y-6=0$, c) $\pi\colon 4x-z-13=0$.

3. Napisać równanie płaszczyzny przechodzącej przez punkt P(2,0,1) i zawierającej prostą l, jeśli:

   a) $l\colon 2x=-y=z+1$, b) $l\colon \begin{cases}x+3y-1=0,\\ 2x-z-8=0,\end{cases}$ c) $l\colon \begin{cases}x=3+2t,\\ y=2,\\ z=4t.\end{cases}$

4. Napisać równanie płaszczyzny przechodzącej przez punkt P(2,0,1) i równoległej do dwu prostych l_1, l_2, jeśli:

   a) $l_1\colon 2x=-y=z+1$, $l_2\colon \begin{cases}x+2y-z=0,\\ x-3z-8=0.\end{cases}$ b) $l_1\colon \begin{cases}x=3+2t,\\ y=2,\\ z=4t.\end{cases}$ $l_2\colon \begin{cases}x+3y-1=0,\\ 2x-z-8=0.\end{cases}$

5. Napisać równanie płaszczyzny przechodzącej przez punkt P(2,0,1), prostopadłej do płaszczyzny π i równoległej do prostej l, jeśli:

   a) $\pi\colon 3x-y-z-10=0$, $l\colon x=3y=z+1$,

   b) $\pi\colon x-2z-13=0$, $l\colon \begin{cases}x+y-3=0,\\ 2x-z-5=0.\end{cases}$

6. Napisać równanie płaszczyzny zawierającej dwie równoległe proste l_1, l_2, jeśli:

   a) $l_1\colon x=y+1=3-2z$, $l_2\colon \begin{cases}x-y-1=0,\\ y+2z=0,\end{cases}$

   b) $l_1\colon \begin{cases}x=2+3t,\\ y=6t,\\ z=2-3t,\end{cases}$ $l_2\colon \begin{cases}x=2-4t,\\ y=-8t,\\ z=1+4t.\end{cases}$

7. Napisać równania kierunkowe prostej przechodzącej przez punkt P(3,4,1), prostopadłej do prostej l i równoległej do płaszczyzny π, jeśli:

   a) $l\colon \frac{x+1}{2}=3y=z-1$, $\pi\colon 3x+3y-6z-11=0$,

   b) $l\colon \begin{cases}x-y-2z-18=0,\\ y-3z-6=0,\end{cases}$ $\pi\colon x+y-3z-7=0$.

8. Znaleźć rzut prostokątny punktu Q(4,5,6) na płaszczyznę π, jeśli:

   a) $\pi\colon 2x-y-z-9=0$, b) π jest płaszczyzną 0xy,

   c) $\pi\colon 2x-3z-3=0$, d) π jest płaszczyzną 0yz.

9. Znaleźć rzut prostokątny punktu Q(4,5,11) na prostą l, jeśli:

   a) $l\colon x+1=3y-1=2-z$, b) $l\colon \begin{cases}x+y-2=0,\\ y+3z-25=0,\end{cases}$

   c) l: jest osią 0x, d) l: jest osią 0y.

10. Znaleźć punkt symetryczny do punktu P(1,2,−3) względem:

    a) płaszczyzny 0xy, b) płaszczyzny 0xz, c) osi 0x, d) osi 0z,

    e) płaszczyzny $\pi\colon x+2y-z+10=0$,

    f) prostej $l\colon \begin{cases}x+2y-z-6=0,\\ x+y+z+3=0,\end{cases}$

    g) prostej $l\colon 5x-15=y-6=5z-10$.

11. Znaleźć rzut prostokątny prostej l na płaszczyznę π, jeśli:
a) $l: x=y+3=2z$, $\pi: 2x+2y+z-3=0$,
b) $l: \begin{cases} x+y+3z-8=0, \\ 2x-y+2=0, \end{cases}$ $\pi: x+2y-z-8=0$.

12. Znaleźć równania krawędziowe prostej l, która jest prostokątnym rzutem prostej l_1 na płaszczyznę π, jeśli:
a) $l_1: x=y+3=2z$, $\pi: 2x+y+4z+13=0$,
b) $l_1: \begin{cases} x+y+3z-8=0, \\ 2x-y-1=0, \end{cases}$ $\pi: x+y-2z+2=0$.

13. Znaleźć równania parametryczne prostej l z zadania 12.

14. Obliczyć odległość prostej $l_1: \frac{x+1}{2}=2-y=3z$ od prostej l_2, jeśli:
a) $l_2: \frac{x+3}{6}=\frac{7-y}{3}=z+19$, b) $l_2: \frac{x-5}{6}=\frac{y+1}{-3}=z-1$,
c) $l_2: \frac{x+1}{2}=\frac{3y}{-2}=3z$, d) $l_2: x=y=z$.

15. Obliczyć odległość punktu P(5,1,3) od prostej l, jeśli:
a) $l: \begin{cases} x+y+z-9=0 \\ 2x-z-7=0 \end{cases}$, b) $l: \frac{x}{2}=4-y=2-z$.

16. Napisać równania parametryczne prostej przechodzącej przez punkt P(-2,3,4) i przecinającej prostą l pod kątem prostym, jeśli:
a) $l: x=y/2-1=3z$, b) $l: 2x+6=y-3=2-z$,
c) $l: \begin{cases} x+3z+6=0, \\ 2x-y+1=0, \end{cases}$ d) $l: \begin{cases} x+y+z-3=0, \\ 2x+y+z-4=0, \end{cases}$
e) l jest osią 0x, f) l jest osią 0z.

Odpowiedzi.

1. a) $18x-6y+z-20=0$, b) $3x+7y-z-1=0$, c) $x-3z+5=0$.
2. a) $x-1=-y/3=2-z$, b) $\begin{cases} x=1+t, \\ y=-3t, \\ z=2, \end{cases}$ c) $\begin{cases} x=1+4t, \\ y=0, \\ z=2-t. \end{cases}$
3. a) $2x-y-2z-2=0$, b) $7x+15y-z-13=0$, c) $4x-3y-2z-6=0$.
4. a) $y+z-1=0$, b) $2x-z-3=0$.

5. a) $x+6y-3z+1=0$, b) $2x+4y+z-5=0$.
6. a) $x-y-1=0$, b) $2x-y-4=0$.
7. a) $x-3=\frac{4-y}{3}=1-z$, b) $\frac{x-3}{5}=\frac{4-y}{8}=1-z$.
8. a) (8,3,4), b) (4,5,0), c) (6,5,3), d) (0,5,6).
9. a) (−41/19,−1/19,60/19), b) (1,1,8), c) (4,0,0), d) (0,5,0).
10. a) (1,2,3), b) (1,-2,-3), c) (1,-2,3), d) (-1,-2,-3), e) (-5,-10,3), f) (-1,0,-5), g) (3,0,5).
11. a) (2,-1,1), b) (1,4,1); rzutem jest punkt, ponieważ $l \perp \pi$.
12. a) $\begin{cases} 2x+y+4z+13=0, \\ 7x-6y-2z-18=0, \end{cases}$ b) $\begin{cases} x+y-2z+2=0, \\ 3x-y+z-4=0. \end{cases}$ Wsk. skorzystać z równania pęku płaszczyzn.
13. a) $\begin{cases} x=-2+22t, \\ y=-5+32t, \\ z=-1-19t, \end{cases}$ b) $\begin{cases} x=1+t, \\ y=1+7t, \\ z=2+4t. \end{cases}$
14. a) $2\sqrt{86}$, b) 0, proste pokrywają się, c) 0, proste przecinają się, d) $\frac{6}{\sqrt{122}}$.
15. a) 0: $P \in l$, b) $\sqrt{11}$.
16. a) $x=-2+52t$, $y=3-11t$, $z=4-90t$, b) $x=-2+2t$, $y=3+t$, $z=4+2t$,
c) $x=-2+t$, $y=3-t$, $z=4-3t$, d) $x=-2-6t$, $y=3+5t$, $z=4+5t$,
e) $x=-2$, $y=3t$, $z=4t$, f) $x=2t$, $y=-3t$, $z=4$.

5. POWIERZCHNIE.

UWAGI OGÓLNE. Jeżeli zbiór punktów P(x,y,z) przestrzeni R^3, których współrzędne x,y,z spełniają równanie

(5.1) $$F(x,y,z)=0,$$

jest pewną powierzchnią, to równanie (5.1) nazywamy równaniem tej powierzchni. Mówimy również, że równanie (5.1) określa (przedstawia) tę powierzchnię. Oczywiście nie zawsze równanie (5.1) określa jakąś powierzchnię.

W szczególności: a) równanie

$$Ax+By+Cz+D=0, \quad \text{gdzie } A^2+B^2+C^2>0$$

jest równaniem płaszczyzny,

b) równanie

$$(x-1)^2+y^2+(z+3)^2=9$$

określa sferę o środku S(1,0,−3) i promieniu 3,

c) równanie

$$(x-1)^2+y^2+(z+3)^2=0$$

równoważne układowi równań $x=1$, $y=0$, $z=-3$ określa punkt P(1,0,-3),

d) równanie

$$(x-y+3z-6)^2+(y+z+8)^2=0$$

jest równoważne układowi równań

$$\begin{cases} x-y+3z-6=0, \\ y+z+8=0 \end{cases}$$

i określa prostą, która jest krawędzią przecięcia płaszczyzn $\pi_1: x-y+3z+6=0$ i $\pi_2: y+z+8=0$.

e) natomiast zbiór punktów spełniających równanie

$$x^2+(z-y)^2+8=0$$

jest zbiorem pustym, gdyż nie istnieje taki punkt P(x,y,z), którego współrzędne spełniałyby to równanie.

Załóżmy, że równanie $F(x,y,z)=0$ jest równaniem pewnej powierzchni S. Wówczas:

1) Powierzchnia S jest **symetryczna względem płaszczyzny** 0xy, wtedy i tylko wtedy, gdy wraz z dowolnym punktem P(x,y,z) należącym do powierzchni S również punkt symetryczny do niego względem płaszczyzny 0xy tzn. $P'(x,y,-z)$ należy do tej powierzchni, czyli dla każdego $P(x,y,z)\in S$ mamy

$$F(x,y,z)=0 \quad \wedge \quad F(x,y,-z)=0.$$

Analogiczna prawidłowość zachodzi przy rozważaniu symetrii powierzchni względem pozostałych płaszczyzn układu współrzędnych 0xyz.

2) Powierzchnia jest S **symetryczna względem osi** 0x, wtedy i tylko wtedy, gdy wraz z dowolnym punktem P(x,y,z) należącym do powierzchni S również punkt symetryczny do niego względem osi 0x tzn. $P'(x,-y,-z)$ należy do tej powierzchni, czyli dla każdego $P(x,y,z)\in S$ mamy

$$F(x,y,z)=0 \quad \wedge \quad F(x,-y,-z)=0.$$

Analogiczna prawidłowość zachodzi w przypadku symetrii względem pozostałych osi układu współrzędnych 0xyz.

3) Powierzchnia S jest **symetryczna względem początku układu współrzędnych** 0xyz wtedy i tylko wtedy, gdy wraz z dowolnym punktem P(x,y,z) należącym do powierzchni S również punkt symetryczny do niego względem początku układu współrzędnych 0xyz tzn. $P'(-x,-y,-z)$ należy do tej powierzchni, czyli dla każdego $P(x,y,z)\in S$ mamy

$$F(x,y,z)=0 \quad \wedge \quad F(-x,-y,-z)=0.$$

Na przykład:

a) powierzchnia o równaniu $x^2+y+z=0$ jest symetryczna względem płaszczyzny 0yz, a nie jest symetryczna względem płaszczyzn 0xz, 0xy, osi 0x, 0y, 0z i początku układu współrzędnych 0xyz,

b) powierzchnia o równaniu $xy+z^2=0$ jest symetryczna względem płaszczyzny 0xy, osi 0z, początku układu współrzędnych 0xyz, a nie jest symetryczna względem płaszczyzn 0xz, 0yz i osi 0x, 0y,

c) natomiast powierzchnia o równaniu $x^3+xy+yz-8=0$ nie jest symetryczna względem: żadnej płaszczyzny, osi i początku układu współrzędnych 0xyz.

Dokładniej zajmiemy się powierzchniami stopnia drugiego.

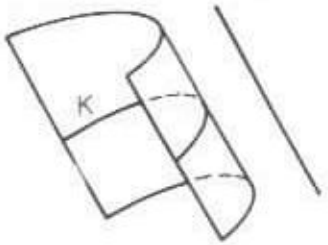
Powierzchnią stopnia drugiego nazywamy zbiór punktów P(x,y,z), których współrzędne x, y, z spełniają równanie stopnia drugiego względem x, y, z

$$A_2x^2+B_2y^2+C_2z^2+A_1xy+B_1xz+C_1yz+Ax+By+Cz+D=0, \tag{5.2}$$

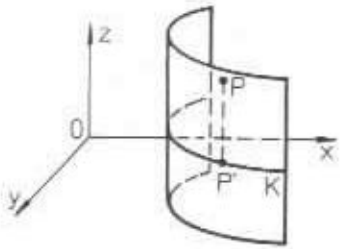
gdzie co najmniej jeden ze współczynników $A_2, B_2, C_2, A_1, B_1, C_1$ jest różny od zera tzn. $A_2^2+B_2^2+C_2^2+A_1^2+B_1^2+C_1^2>0$.

Spośród wszystkich powierzchni stopnia drugiego omówimy kolejno: powierzchnie walcowe, powierzchnię kulistą czyli sferę, elipsoidę, paraboloidy, hiperboloidy i stożek eliptyczny.

POWIERZCHNIE WALCOWE. Powierzchnią walcową nazywamy zbiór prostych równoległych do danej prostej i przechodzących przez punkty krzywej K, przy czym krzywą K nazywamy **kierownicą**, a proste - **tworzącymi** powierzchni walcowej (rys 5.1).



Rys 5.1.



Rys 5.2.

Niech kierownica K powierzchni walcowej W będzie krzywą płaską zawartą w płaszczyźnie 0xy i niech w układzie współrzędnych 0xy kierownica K ma równanie $F(x,y)=0$, czyli

$$K = \{(x,y) \in R^2 : F(x,y)=0\}.$$

Kierownica K w układzie współrzędnych 0xyz opisana jest równaniami

$$K = \{(x,y,z) \in R^3 : F(x,y)=0,\ z=0\}.$$

Tworzącymi takiej powierzchni walcowej są proste równoległe do osi 0z (rys.5.2). Wówczas dowolny punkt P(x,y,z) należy do powierzchni walcowej W wtedy i tylko wtedy, gdy prostokątny rzut P' punktu P na płaszczyznę 0xy należy do krzywej K tzn.

$$P(x,y,z) \in W \Leftrightarrow P'(x,y) \in K \Leftrightarrow F(x,y)=0.$$

Zatem równanie

(5.3) $$F(x,y)=0,$$

jest w przestrzeni R^3 równaniem rozważanej powierzchni walcowej (rys 5.2).

Analogicznie: równanie

$$F(y,z)=0,$$

jest w przestrzeni R^3 równaniem powierzchni walcowej, której kierownicą jest krzywa o równaniu $F(y,z)=0$ na płaszczyźnie 0yz, a tworzące są równoległe do osi 0x. Równanie

$$F(x,z)=0,$$

jest w przestrzeni R^3 równaniem powierzchni walcowej o tworzących równoległych do osi 0y i kierownicy o równaniu $F(x,z)=0$ na płaszczyźnie 0xz.

Na przykład: a) równanie $x^2+z^2=4$ przedstawia na płaszczyźnie 0xz okrąg K o środku w punkcie (0,0) i promieniu r =2, zaś w układzie współrzędnych 0xyz równanie to określa walec kołowy o tworzących równoległych do osi 0y i kierownicy K,

b) równanie $2x+y-3=0$ przedstawia na płaszczyźnie 0xy prostą K, zaś w układzie współrzędnych 0xyz płaszczyznę równoległą do osi 0z (powierzchnię walcową o kierownicy K i tworzących równoległych do osi 0z).

Powierzchnie walcowe, których kierownicami są krzywe stożkowe - elipsa (w szczególności okrąg), hiperbola, parabola - leżące w jednej z płaszczyzn układu współrzędnych, a tworzące są równoległe do odpowiedniej osi układu współrzędnych są powierzchniami stopnia drugiego.

1) Równanie

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1, \quad \text{gdzie } a>0,\ b>0$$

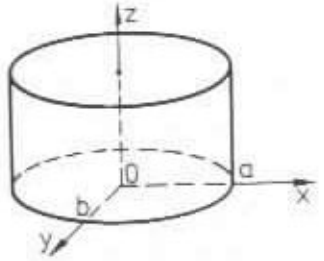
jest w układzie współrzędnych 0xyz równaniem powierzchni walcowej, której kierownicą jest elipsa o równaniu $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ leżąca na płaszczyźnie 0xy, zaś tworzące są równoległe do osi 0z. Powierzchnię tę nazywamy **walcem eliptycznym** (rys 5.3)

Walec eliptyczny jest powierzchnią symetryczną względem wszystkich płaszczyzn, osi i początku układu współrzędnych 0xyz. Przekrój tego walca dowolną płaszczyzną prostopadłą do osi 0z jest elipsą, zaś przekrój płaszczyzną prostopadłą do osi 0x lub 0y jest dwiema prostymi równoległymi (w szczególności jedną prostą lub zbiorem pustym).

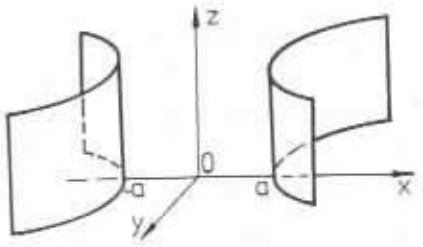
W szczególności, gdy $a=b=r$, otrzymujemy równanie

$$x^2+y^2=r^2,$$

które jest równaniem **walca kołowego**.



Rys 5.3.



Rys 5.4.

2) Równanie

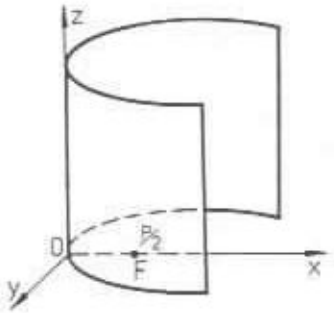
$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1, \quad \text{gdzie } a>0,\ b>0$$

jest w układzie współrzędnych 0xyz równaniem powierzchni walcowej, której kierownicą jest hiperbola o równaniu $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$ leżąca na płaszczyźnie 0xy, zaś tworzące są równoległe do osi 0z. Powierzchnię tę nazywamy **walcem hiperbolicznym** (rys 5.4). Walec ten jest symetryczny względem wszystkich płaszczyzn, osi i początku układu współrzędnych 0xyz.

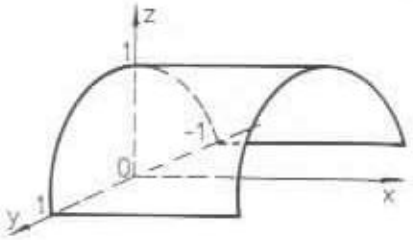
3) Równanie

$$y^2=2px, \quad \text{gdzie } p\neq 0$$

w układzie 0xyz jest równaniem **walca parabolicznego** (rys 5.5 dla $p>0$), którego tworzące są równoległe do osi 0z, zaś kierownicą jest parabola o równaniu $y^2=2px$ na płaszczyźnie 0xy. Walec ten jest symetryczny względem płaszczyzn 0xz, 0xy i osi 0x.



Rys 5.5.



Rys 5.6.

Na przykład równanie $z=1-y^2$ w przestrzeni R^3z układem współrzędnych 0xyz jest równaniem walca parabolicznego o tworzących równoległych do osi 0x i kierownicy $K: z=1-y^2$ na płaszczyźnie 0yz (rys 5.6).

PRZYKŁAD 5.1. Napiszemy równania przekrojów walca eliptycznego o równaniu $x^2+4y^2=100$ płaszczyznami π_1: $x=6$, π_2: $z=4$.

Przekrój płaszczyzną π_1 określają równania

$$\begin{cases} x^2+4y^2=100, \\ x=6, \end{cases} \Leftrightarrow \begin{cases} 36+4y^2=100, \\ x=6, \end{cases} \Leftrightarrow \begin{cases} y=\pm 4, \\ x=6. \end{cases}$$

Przekrojem są więc dwie proste

$$l_1: \begin{cases} y=4, \\ x=6, \end{cases} \qquad l_2: \begin{cases} y=-4, \\ x=6. \end{cases}$$

Przekrój płaszczyzną π_2 jest elipsą o równaniach

$$\begin{cases} x^2+4y^2=100, \\ z=4, \end{cases} \Leftrightarrow \begin{cases} \frac{x^2}{100}+\frac{y^2}{25}=1, \\ z=4. \end{cases}$$ ■

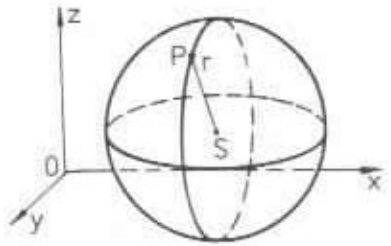
SFERA. Powierzchnią kulistą (sferą) o środku $S(x_0,y_0,z_0)$ i promieniu r, $r>0$, nazywamy zbiór punktów P(x,y,z), których odległość od środka S jest równa r (rys 5.7). Zatem

$$P(x,y,z)\in \text{sfery} \Leftrightarrow |PS|=r \Leftrightarrow \sqrt{(x-x_0)^2+(y-y_0)^2+(z-z_0)^2}=r,$$

skąd otrzymujemy

(5.4) $$(x-x_0)^2+(y-y_0)^2+(z-z_0)^2=r^2, \text{ gdzie } r>0.$$

To ostatnie równanie jest równaniem powierzchni kulistej o środku $S(x_0,y_0,z_0)$ i promieniu r.



Rys 5.7.

Na przykład równanie powierzchni kulistej o środku $S(-1,2,3)$ i promieniu $r=2$ ma postać

$$(x+1)^2+(y-2)^2+(z-3)^2=4,$$

a po wykonaniu działań

$$x^2+y^2+z^2+2x-4y-6z+10=0.$$

W szczególności, gdy $x_0 = y_0 = z_0 = 0$ otrzymujemy równanie

$$x^2 + y^2 + z^2 = r^2.$$

Jest to równanie sfery o środku S(0,0,0), symetrycznej względem wszystkich płaszczyzn, osi i początku układu współrzędnych 0xyz.

Każdy przekrój sfery płaszczyzną jest o k r ę g i e m (w szczególności punktem lub zbiorem pustym).

P R Z Y K Ł A D 5.2. Znajdziemy środek i promień sfery o równaniu

$$x^2 + y^2 + z^2 - 2x + 4y - 4 = 0.$$

Równanie to równoważne jest kolejno równaniom:

$$x^2 - 2x + y^2 + 4y + z^2 - 4 = 0,$$

$$(x-1)^2 - 1 + (y+2)^2 - 4 + z^2 - 4 = 0,$$

$$(x-1)^2 + (y+2)^2 + z^2 = 9.$$

Dane równanie przedstawia sferę o środku S(1,-2,0) i promieniu $r = 3$. ■

P R Z Y K Ł A D 5.3. Pokażemy, że równanie

$$z = 2 - \sqrt{16 - x^2 - y^2}$$

jest równaniem "dolnej" połowy sfery o równaniu $x^2 + y^2 + (z-2)^2 = 16$.

Stosując przekształcenia równoważne otrzymujemy

$$z = 2 - \sqrt{16 - x^2 - y^2} \Leftrightarrow z - 2 = -\sqrt{16 - x^2 - y^2} \Leftrightarrow$$

$$\Leftrightarrow [(z-2)^2 = 16 - x^2 - y^2 \wedge z - 2 \le 0] \Leftrightarrow$$

$$\Leftrightarrow [x^2 + y^2 + (z-2)^2 = 16 \wedge z \le 2].$$

Równanie $x^2 + y^2 + (z-2)^2 = 16$ jest równaniem sfery o środku S(0,0,2) i promieniu $r = 4$. Zatem równanie $z = 2 - \sqrt{16 - x^2 - y^2}$ jest równaniem części tej sfery dla $z \le 2$, czyli równaniem "dolnej" połowy tej sfery. ■

ELIPSOIDA. Powierzchnię o równaniu

$$(5.5) \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad \text{gdzie } a>0,\ b>0,\ c>0,$$

nazywamy **elipsoidą**.

Łatwo sprawdzić, że jest to powierzchnia symetryczna względem wszystkich płaszczyzn, osi i początku układu współrzędnych 0xyz. Powierzchnia ta przecina osie 0x, 0y, 0z odpowiednio w punktach $(\pm a,0,0)$, $(0,\pm b,0)$, $(0,0,\pm c)$ zwanych **wierzchołkami** tej elipsoidy. Kształt powierzchni badamy analizując przekroje płaszczyznami prostopadłymi do osi układu współrzędnych 0xyz. Przekrój płaszczyzną prostopadłą do osi 0z ma równania

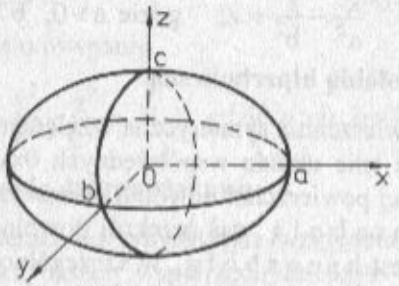
$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \\ z = k, \end{cases} \Leftrightarrow \begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 - \frac{k^2}{c^2}, \\ z = k. \end{cases}$$

Jeśli $1 - \frac{k^2}{c^2} > 0$, czyli $|k| < c$, to przekrój jest e l i p s ą. "Największą" spośród tych elips otrzymujemy dla $k = 0$ i jest to elipsa o równaniach

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \\ z = 0, \end{cases}$$

której półosiami są a i b. Analogicznie otrzymujemy, że: przekrój płaszczyzną $x = k$, gdy $|k| < a$ lub $y = k$, gdy $|k| < b$ jest e l i p s ą.

Elipsoidę o równaniu (5.5) przestawia rysunek 5.8.



Rys 5.8.

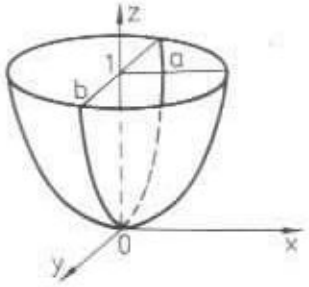
PARABOLOIDY. Powierzchnię o równaniu

$$(5.6) \qquad \frac{x^2}{a^2} + \frac{y^2}{b^2} = z, \quad \text{gdzie } a>0,\ b>0,$$

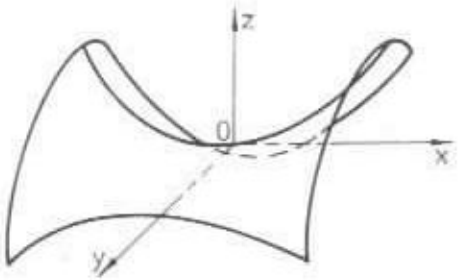
nazywamy **paraboloidą eliptyczną**.

Jest to powierzchnia symetryczna względem płaszczyzn 0yz, 0xz i osi 0z; przecina osie układu współrzędnych 0xyz jedynie w punkcie (0,0,0) zwanym **wierzchołkiem** tej paraboloidy. Przekrój tej powierzchni dowolną płaszczyzną prostopadłą do osi 0x lub 0y jest p a r a b o l ą , zaś przekrój płaszczyzną prostopadłą do osi 0z o równaniu $z = k$, gdy $k > 0$, jest e l i p s ą . Paraboloidę eliptyczną o równaniu (5.6) przedstawia rysunek 5.9.

W szczególności, gdy $a = b$, powierzchnię (5.6) nazywamy paraboloida obrotową.



Rys 5.9.



Rys 5.10.

Powierzchnię o równaniu

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z, \quad \text{gdzie } a>0,\ b>0, \tag{5.7}$$

nazywamy **paraboloidą hiperboliczną**.

Jest to powierzchnia symetryczna względem płaszczyzn 0xz, 0yz i osi 0z; przecina osie układu współrzędnych 0xyz jedynie w punkcie (0,0,0). Przekrój tej powierzchni dowolną płaszczyzną prostopadłą do osi 0x lub 0y jest p a r a b o l ą , zaś przekrój dowolną płaszczyzną prostopadłą do osi 0z jest h i p e r b o l ą , w szczególności dwiema prostymi. Paraboloidę hiperboliczną o równaniu (5.7) przedstawia rysunek 5.10.

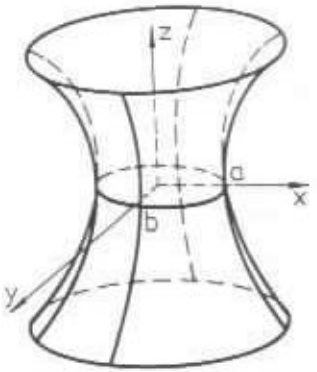
HIPERBOLOIDY. Powierzchnię o równaniu

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \quad \text{gdzie } a>0,\ b>0,\ c>0, \tag{5.8}$$

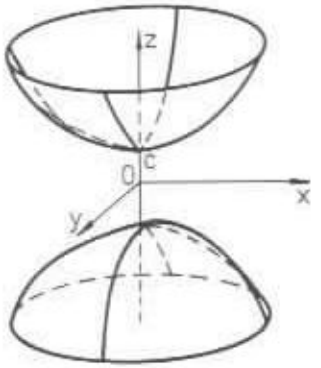
nazywamy **hiperboloidą jednopowłokową**.

Jest to powierzchnia symetryczna względem wszystkich płaszczyzn, osi i początku układu współrzędnych 0xyz, przecina osie 0x i 0y odpowiednio w punktach $(\pm a,0,0)$ i $(0,\pm b,0)$. Przekrój tej powierzchni dowolną płaszczyzną prostopadłą do osi 0x lub 0y jest h i p e r b o l ą , w szczególności dwiema prostymi, zaś przekrój dowolną płaszczyzną prostopadłą do osi 0z jest e l i p s ą . Hiperboloidę jednopowłokową o równaniu (5.8) przedstawia rysunek 5.11.

W szczególności, gdy $a = b$, otrzymujemy hiperboloidę jednopowłokową obrotową.



Rys 5.11.



Rys 5.12.

Powierzchnię o równaniu

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1, \quad \text{gdzie } a>0,\ b>0,\ c>0, \tag{5.9}$$

nazywamy **hiperboloidą dwupowłokową**.

Jest to powierzchnia symetryczna względem wszystkich płaszczyzn, osi i początku układu współrzędnych 0xyz. Przecina oś 0z w punktach $(0,0,\pm c)$. Przekrój tej powierzchni dowolną płaszczyzną prostopadłą do osi 0x lub 0y jest h i p e r b o l ą , zaś przekrój płaszczyzną prostopadłą do osi 0z o równaniu $z = k$, gdy $|k| > c$, jest e l i p s ą . Hiperboloidę dwupowłokową o równaniu (5.9) przedstawia rysunek 5.12.

W szczególności, gdy $a = b$, otrzymujemy hiperboloidę dwupowłokową obrotową.

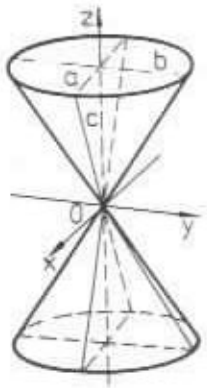
STOŻEK ELIPTYCZNY. Powierzchnię o równaniu

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=0, \quad \text{gdzie } a>0,\ b>0,\ c>0, \tag{5.10}$$

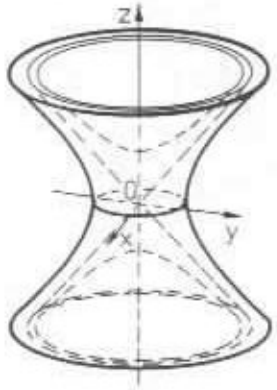
nazywamy **stożkiem eliptycznym** (stożkiem asymptotycznym lub krótko: stożkiem).

Jest to powierzchnia symetryczna względem wszystkich płaszczyzn, osi i początku układu współrzędnych 0xyz; przecina osie układu jedynie w punkcie (0,0,0) zwanym **wierzchołkiem** tego stożka. Przekrój tej powierzchni płaszczyzną $x=k$ lub $y=k$, gdy $k\neq 0$, jest h i p e r b o l ą, a dla $k=0$ - dwiema prostymi, zaś przekrój płaszczyzną $z=k$, gdy $k\neq 0$ jest e l i p s ą, a dla $k=0$ jest punktem (0,0,0). Stożek eliptyczny o równaniu (5.10) przedstawia rysunek 5.13.

W szczególności, gdy $a=b$, stożek nazywamy obrotowym.



Rys 5.13.



Rys 5.14.

Rysunek 5.14 przedstawia hiperboloidę jedno-, dwupowłokową i stożek eliptyczny o równaniach (5.8) - (5.10). Rysunek ten wyjaśnia dlaczego stożek eliptyczny nazywa się również **stożkiem asymptotycznym**.

Na koniec tego skrótowego przeglądu powierzchni stopnia drugiego odnotujmy, że wśród wymienionych powierzchni są powierzchnie:

1) **obrotowe** tzn. takie, które powstają przez obrót krzywej dookoła pewnej prostej,

2) **prostokreślne** tzn. takie, że przez każdy punkt powierzchni można poprowadzić prostą zawartą w tej powierzchni.

Powierzchniami obrotowymi są: walec obrotowy, sfera, elipsoida (5.5) dla $a=b$ lub $a=c$ lub $b=c$ oraz paraboloida eliptyczna (5.6), hiperboloidy (5.8) i (5.9) i stożek eliptyczny (5.10) dla $a=b$.

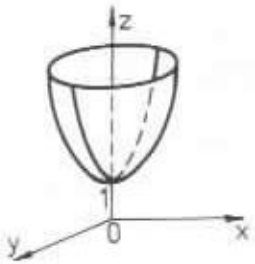
Powierzchniami prostokreślnymi są: oczywiście wszystkie powierzchnie walcowe, stożek eliptyczny oraz hiperboloida jednopowłokowa i paraboloida hiperboliczna.

P R Z Y K Ł A D 5.4. a) Równanie $z=1+x^2+y^2$ jest równaniem paraboloidy obrotowej (rys 5.15) o wierzchołku (0,0,1).

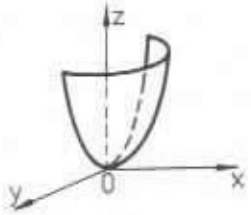
b) Równanie $x=\sqrt{z-y^2}$ przedstawia pewną powierzchnię. Dla $x\geq 0$ mamy

$$x=\sqrt{z-y^2} \Leftrightarrow x^2=z-y^2 \Leftrightarrow z=x^2+y^2.$$

To ostatnie równanie określa paraboloidę obrotową, zatem równanie $x=\sqrt{z-y^2}$ określa tę część paraboloidy $z=x^2+y^2$, która jest zawarta w półprzestrzeni $x\geq 0$ (rys 5.16). ■



Rys 5.15.



Rys 5.16.

ZADANIA DO ROZWIĄZANIA.

1. Określić, które z powierzchni o niżej podanych równaniach są symetryczne względem płaszczyzn, osi lub początku układu współrzędnych 0xyz:

a) $x^2-y^2-xz=0$, b) $y=z^2$, c) $x^2-3y^2-z^2=6$,
d) $x^2+y^2-3z=0$, e) $xyz=3$, f) $x^3+y-z=6$.

2. Wyznaczyć zbiór punktów $P(x,y,z) \in R^3$ określony równaniem:
   a) $x^2+y^2=0$, b) $x^2+(y-1)^2+z^2=-4$,
   c) $(x-y)(x+z)=0$, d) $(x-y)^2+(x+z)^2=0$.

3. Nazwać i naszkicować zbiór punktów $P(x,y) \in R^2$ określony równaniem:
   a) $x^2-y^2=1$, b) $x^2+4y^2=1$, c) $y^2-x=0$.

4. Nazwać i naszkicować zbiór punktów $P(x,y,z) \in R^3$ określony równaniem:
   a) $x^2-y^2=1$, b) $x^2+4y^2=1$, c) $z^2-x=0$,
   d) $y^2+(z-2)^2=4$, e) $z^2+y^2=4$, f) $yz=4$.

5. Nazwać i naszkicować powierzchnię o równaniu:
   a) $3x^2+y^2+z^2=48$, b) $z=\sqrt{4-x^2-4y^2}$, c) $z=\sqrt{x^2+4y^2-4}$,
   d) $z=(x-1)^2+y^2$, e) $z=\sqrt{(x-1)^2+y^2}$, f) $z=\sqrt{(x-1)^2-y^2}$.

Odpowiedzi.

1. a) Symetria wzgl. 0xz, 0y, (0,0,0), b) 0xy, 0yz, 0y, c) symetria wzgl. wszystkich płaszczyzn, osi i początku układu współrzędnych 0xyz, d) 0yz, 0xz, 0z, e) 0x, 0y, 0z, f) brak symetrii.
2. a) Oś 0z; $x^2+y^2=0 \Leftrightarrow x=0, y=0$ i z dowolne, b) zbiór pusty, c) suma $\pi_1 \cup \pi_2$ płaszczyzn: $\pi_1: x-y=0$, $\pi_2: x+z=0$, d) iloczyn $\pi_1 \cap \pi_2$ płaszczyzn z c).
3. a) Hiperbola, b) elipsa, c) parabola.
4. a) Walec hiperboliczny, tworzące równoległe do osi 0z, b) walec eliptyczny, tworzące równoległe do osi 0z, c) walec paraboliczny, tworzące równoległe do 0y, d) walec kołowy, tworzące równoległe do osi 0x, e) walec kołowy, tworzące równoległe do osi 0x, f) walec hiperboliczny, tworzące równoległe do osi 0x.
5. a) Elipsoida obrotowa, b) "górna" połowa elipsoidy; $z \geq 0$, c) "górna" połowa hiperboloidy jednopowłokowej; $z \geq 0$, d) paraboloida obrotowa o wierzchołku (1,0,0), e) "górna" połowa stożka obrotowego o wierzchołku (1,0,0), ułożonego wzdłuż osi 0z, f) "górna" połowa stożka obrotowego o wierzchołku (1,0,0) ułożonego wzdłuż osi 0x.

II. RACHUNEK RÓŻNICZKOWY FUNKCJI WIELU ZMIENNYCH

1. PRZESTRZEŃ METRYCZNA I PEWNE WŁASNOŚCI JEJ PODZBIORÓW.

PRZESTRZEŃ METRYCZNA. Pojęcie przestrzeni metrycznej zostało już wprowadzone w pierwszym tomie tej książki (rozdz. I, 4). O ile w dotychczasowych rozważaniach można było je pominąć, to obecnie, przy omawianiu funkcji wielu zmiennych, pojęcie przestrzeni metrycznej jest bardzo przydatne i nie warto z niego rezygnować. Dlatego też proponujemy, aby Czytelnik zechciał jeszcze raz przeczytać wskazany paragraf, mimo że pewne wiadomości dotyczące tego zagadnienia zostaną tu powtórzone.

Załóżmy, że dany jest niepusty zbiór X i funkcja ρ, która każdej parze elementów zbioru X przyporządkowuje liczbę rzeczywistą, przy czym funkcja ta dla dowolnych $p_1, p_2, p_3 \in X$ spełnia następujące warunki:

(1) $\rho(p_1,p_2)=0 \Leftrightarrow p_1=p_2$,

(2) $\rho(p_1,p_2)=\rho(p_2,p_1)$,

(3) $\rho(p_1,p_2) \leq \rho(p_1,p_3)+\rho(p_3,p_2)$.

Funkcję ρ nazywamy **metryką** lub **odległością** w zbiorze X, a zbiór X z metryką ρ nazywamy **przestrzenią metryczną** i oznaczamy symbolem (X,ρ). Liczbę $\rho(p_1,p_2)$ nazywamy **odległością elementów** p_1 i p_2, a warunki (1), (2), (3) - aksjomatami metryki.

Z warunków (1), (2) i (3) wynika, że wartości funkcji ρ są liczbami nieujemnymi. Istotnie, dla dowolnych $p_1, p_2 \in X$ mamy

$$0=\rho(p_1,p_1)\le\rho(p_1,p_2)+\rho(p_2,p_1)=2\rho(p_1,p_2),$$

skąd wynika, że

$$\rho(p_1,p_2)\ge 0.$$

Przypomnijmy jeszcze, że: a) w dowolnym niepustym zbiorze można określić metrykę, co oznacza, że każdy niepusty zbiór można "zmetryzować", b) w ustalonym niepustym zbiorze można określić metrykę na różne sposoby (por. tom 1, rozdz. I, 4).

W przypadku gdy metryka w X jest ustalona i nie może być wątpliwości z jaką metryką zbiór X jest rozważany, albo też wiadomo, że w zbiorze X jest określona metryka i nie jest ważne jaka to jest metryka, pisać będziemy krócej: przestrzeń metryczna X.

PRZESTRZEŃ EUKLIDESOWA R^n. Zbiór wszystkich n-wyrazowych ciągów liczb rzeczywistych, tzn. zbiór $X=\underbrace{R\times\cdots\times R}_{n}=R^n$ z metryką ρ określoną wzorem

$$\rho(p_1,p_2)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}, \tag{1.1}$$

gdzie $p_1=(x_1,x_2,\ldots,x_n)$, $p_2=(y_1,y_2,\ldots,y_n)$ są dowolnymi elementami (punktami) zbioru X, nazywamy **n-wymiarową przestrzenią euklidesową** i oznaczamy krótko symbolem R^n. Metrykę (1.1) nazywamy **metryką euklidesową** lub **naturalną**.

W szczególności: przestrzeń $R^1\equiv R$ oznacza zbiór liczb rzeczywistych x, w którym odległość dowolnych dwóch elementów $p_1=x_1$ i $p_2=x_2$ określa wzór:

$$\rho(p_1,p_2)=\rho(x_1,x_2)=\sqrt{(x_1-x_2)^2}=|x_1-x_2|;$$

Podobnie R^2 oznacza zbiór par liczb rzeczywistych (x,y), w którym odległość dowolnych dwóch elementów $p_1=(x_1,y_1)$ i $p_2=(x_2,y_2)$ określa wzór:

$$\rho(p_1,p_2)=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2};$$

Przestrzeń R^3 oznacza zbiór uporządkowanych trójek liczb rzeczywistych (x,y,z), w którym odległość dowolnych dwóch elementów $p_1=(x_1,y_1,z_1)$ i $p_2=(x_2,y_2,z_2)$ określona jest wzorem:

$$\rho(p_1,p_2)=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}.$$

Obrazem geometrycznym przestrzeni R^1 jest oś liczbowa, przestrzeni R^2 - płaszczyzna z kartezjańskim układem współrzędnych 0xy, a R^3 - przestrzeń z kartezjańskim układem współrzędnych 0xyz, przy czym odległość dowolnych dwóch punktów każdego z tych zbiorów mierzona jest długością odcinka łączącego te punkty. Dla przestrzeni R^n, gdy $n>3$, nie ma już interpretacji geometrycznej.

KULA. ZBIORY OGRANICZONE. Niech (X,ρ) będzie dowolną przestrzenią metryczną. Uogólniając znaną nam definicję kuli przyjmijmy następujące określenie.

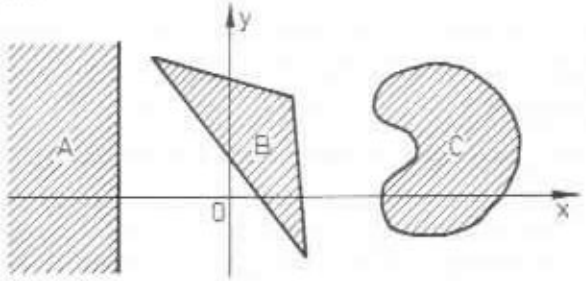
W przestrzeni metrycznej (X,ρ) **kulą otwartą** (krótko: **kulą**) o środku w punkcie $a\in X$ i promieniu r, $r>0$, nazywamy zbiór

$$K(a,r)=\{p\in X:\ \rho(p,a)<r\}.$$

Kula $K(a,r)$ w przestrzeni R^1 oznacza więc otwarty przedział o środku w punkcie a i długości 2r, w przestrzeni R^2 - wnętrze koła o środku w punkcie a i promieniu r, a w R^3 - wnętrze kuli o środku w punkcie a i promieniu r.

Zbiór $A\subset X$ nazywamy **ograniczonym**, gdy istnieje pewna kula $K(a,r)\subset X$ taka, że $A\subset K(a,r)$.

Zbiór nazywamy **nieograniczonym**, gdy nie jest on zbiorem ograniczonym.



Rys 1.1

Na rysunku 1.1 pokazane są pewne podzbiory przestrzeni R^2. Zbiór A jest nieograniczony, a zbiory B i C są ograniczone.

ZBIORY OTWARTE I DOMKNIĘTE. Niech (X,ρ) będzie dowolną przestrzenią metryczną. Przyjmujemy następujące definicje.

Otoczenie punktu p_0 o promieniu ε, $\varepsilon>0$, oznaczane symbolem $U(p_0,\varepsilon)$, jest to kula o środku w punkcie p_0 i promieniu ε. Zatem $U(p_0,\varepsilon)=K(p_0,\varepsilon)$, czyli

$$U(p_0,\varepsilon)=\{p\in X:\ \rho(p,p_0)<\varepsilon\}.$$

Sąsiedztwo punktu p_0 o promieniu ε, $\varepsilon>0$, jest to zbiór $S(p_0,\varepsilon)=U(p_0,\varepsilon)-\{p_0\}$, czyli

$$S(p_0,\varepsilon)=\{p\in X:\ 0<\rho(p,p_0)<\varepsilon\}.$$

W przestrzeni R^1 mamy więc:

$$U(x_0,\varepsilon)=\{x\in R:\ |x-x_0|<\varepsilon\}=(x_0-\varepsilon,x_0+\varepsilon),$$

$$S(x_0,\varepsilon)=\{x\in R:\ 0<|x-x_0|<\varepsilon\}=(x_0-\varepsilon,x_0)\cup(x_0,x_0+\varepsilon)$$

i jest to zgodne z wcześniej przyjętymi określeniami otoczenia i sąsiedztwa punktu w zbiorze R. Interpretację geometryczną otoczenia i sąsiedztwa punktu w R^2 i R^3 pozostawiamy Czytelnikowi.

Jeżeli promień otoczenia lub sąsiedztwa punktu p_0 nie jest istotny w naszych rozważaniach, będziemy używać krótszego oznaczenia: $U(p_0)$ lub $S(p_0)$.

Punkt $p_0\in A\subset X$ nazywamy **punktem wewnętrznym zbioru** A, gdy zbiór A zawiera pewne otoczenie tego punktu, to znaczy: istnieje otoczenie $U(p_0)$ takie, że $U(p_0)\subset A$.

Punkt p_0 nazywamy **punktem zewnętrznym zbioru** $A\subset X$, gdy istnieje otoczenie $U(p_0)$ tego punktu, do którego nie należy żaden punkt zbioru A, to znaczy, że $U(p_0)\cap A=\emptyset$.

Punkt p_0 nazywamy **punktem brzegowym zbioru** $A\subset X$, gdy każde otoczenie tego punktu zawiera co najmniej jeden punkt należący do zbioru A i co najmniej jeden punkt do tego zbioru nie należący.

Zbiór wszystkich punktów brzegowych zbioru nazywamy **brzegiem** tego zbioru.

Punkt p_0 nazywamy **punktem skupienia zbioru** $A\subset X$, gdy każde otoczenie tego punktu zawiera co najmniej jeden, różny od p_0, punkt zbioru A.

Punkt p_0 należący do zbioru A, który nie jest jego punktem skupienia, nazywamy **punktem izolowanym** lub **odosobnionym** tego zbioru.

Z podanych określeń wynika, że punkt brzegowy oraz punkt skupienia zbioru mogą, ale nie muszą do tego zbioru należeć.

PRZYKŁAD 1.1.

a) W przestrzeni R^2 weźmy pod uwagę odcinek osi odciętych:

$$A=\{(x,y)\in R^2:\ -1\le x\le 1\ \wedge\ y=0\}.$$

Z przyjętych wcześniej definicji wynika, że każdy punkt zbioru A jest punktem brzegowym i jednocześnie punktem skupienia tego zbioru. Zbiór A nie ma punktów wewnętrznych.

Zauważmy, że ten sam odcinek może być traktowany jako podzbiór przestrzeni R^1:

$$B=\{x\in R:\ -1\le x\le 1\}.$$

Wówczas każdy jego punkt, z wyjątkiem $x=-1$ oraz $x=1$, jest punktem wewnętrznym zbioru B, a punkty $x=-1$ i $x=1$ są punktami brzegowymi tego zbioru. Każdy punkt zbioru B jest punktem skupienia tego zbioru.

b) Niech A oznacza dowolne koło na płaszczyźnie bez okręgu tego koła. Każdy punkt zbioru A jest punktem wewnętrznym tego zbioru. Każdy punkt leżący na okręgu koła jest punktem brzegowym, a każdy punkt koła A i każdy punkt leżący na okręgu tego koła jest punktem skupienia zbioru A.

c) Załóżmy, że A jest podzbiorem przestrzeni R^2 złożonym z dowolnej, ale skończonej liczby punktów. Taki zbiór nie ma punktów skupienia, ani też punktów wewnętrznych. Każdy punkt takiego zbioru jest punktem odosobnionym, a także punktem brzegowym tego zbioru. ■

Zbiór $A\subset X$ nazywamy **otwartym** w przestrzeni X, gdy każdy punkt tego zbioru jest jego punktem wewnętrznym.

Zbiór $A \subset X$ nazywamy **domkniętym** w przestrzeni X, gdy zawiera on wszystkie swoje punkty skupienia.

Największy zbiór otwarty zawarty w A nazywamy **wnętrzem zbioru** A i oznaczamy symbolem Int A.

Najmniejszy zbiór domknięty zawierający zbiór A nazywamy **domknięciem zbioru** A i oznaczamy symbolem $\overline{A}$.

Z określeń tych łatwo wynika, że $\operatorname{Int} A = A$ jedynie wtedy, gdy A jest zbiorem otwartym oraz $A = \overline{A}$ jedynie wtedy, gdy A jest zbiorem domkniętym.

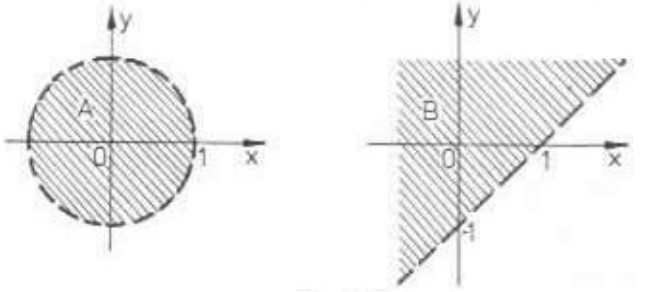
PRZYKŁAD 1.2. Zbiory

$$A = \{(x,y) \in R^2 : \ x^2 + y^2 < 1\}, \qquad B = \{(x,y) \in R^2 : \ y > x - 1\}$$

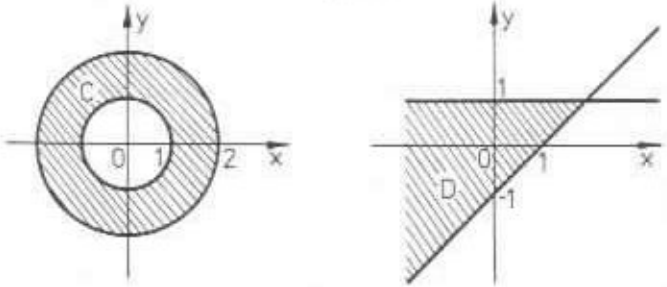
są zbiorami otwartymi na płaszczyźnie, natomiast zbiory

$$C = \{(x,y) \in R^2 : 1 \le x^2 + y^2 \le 4\}, \qquad D = \{(x,y) \in R^2 : \ x - 1 \le y \le 1\}$$

są zbiorami domkniętymi na płaszczyźnie (rys 1.2 i 1.3).



Rys 1.2



Rys 1.3

Zbiorem domkniętym jest również każdy zbiór skończony (złożony ze skończonej liczby elementów), gdyż taki zbiór nie ma punktów skupienia.

Natomiast zbiór

$$E = \{(x,y) \in R^2 : \ 1 \le x^2 + y^2 < 4\}$$

nie jest zbiorem otwartym (gdyż punkty leżące na okręgu $x^2 + y^2 = 1$ nie są punktami wewnętrznymi tego zbioru) i nie jest zbiorem domkniętym (gdyż punkty leżące na okręgu $x^2 + y^2 = 4$ są punktami skupienia tego zbioru, ale do niego nie należą). ■

PRZYKŁAD 1.3.

a) Niech $A = \{x \in R : \ 1 < x \le 2\}$. Wówczas

$$\operatorname{Int} A = \{x \in R : \ 1 < x < 2\}, \qquad \overline{A} = \{x \in R : \ 1 \le x \le 2\}.$$

b) Niech $A = \{(x,y) \in R^2 : \ x^2 + y^2 < 1\}$. Wówczas

$$\operatorname{Int} A = A, \qquad \overline{A} = \{(x,y) \in R^2 : \ x^2 + y^2 \le 1\}.$$

c) Niech $A = \{(x,y) \in R^2 : \ 1 \le x \le 2 \wedge -1 \le y \le 3\}$. Wówczas

$$\operatorname{Int} A = \{(x,y) \in R^2 : \ 1 < x < 2 \wedge -1 < y < 3\}, \qquad \overline{A} = A.$$

Czytelnikowi pozostawiamy interpretację geometryczną powyższych przykładów. ■

OBSZARY. Niech X będzie dowolną przestrzenią metryczną.

Zbiór $A \subset X$ nazywamy **spójnym**, gdy przy każdym przedstawieniu tego zbioru jako sumy dwóch zbiorów r o z ł ą c z n y c h i n i e p u s t y c h A_1 i A_2, przynajmniej jeden z tych zbiorów zawiera punkty skupienia drugiego zbioru.

U w a g a . W przestrzeniach R^2 i R^3 przyjmuje się także zbiór nazywać spójnym, gdy dowolne dwa punkty tego zbioru można połączyć linią łamaną zawartą w tym zbiorze.

Zbiór spójny i otwarty nazywamy **obszarem**. Domknięcie obszaru nazywamy **obszarem domkniętym**.

Zgodnie z przyjętym określeniem każdy przedział otwarty w przestrzeni R^1 jest obszarem. Na przykład zbiory

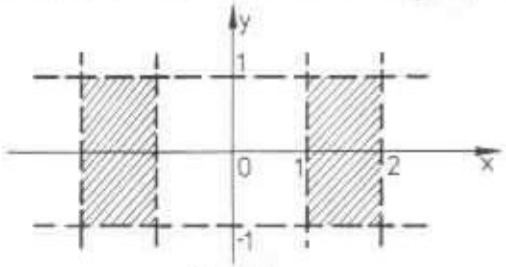
$$A=\{x\in R:\ a<x<b\},\ a<b,\qquad B=\{x\in R:\ a<x<+\infty\}$$

są obszarami. Jednakże w przypadku przestrzeni R^1 termin ten nie jest używany. W dalszym ciągu będziemy mówić o obszarach jedynie w przestrzeniach R^n, gdy $n\geq 2$.

W przestrzeni R^2 obszarami, na przykład, są: wnętrze koła, elipsy, dowolnego trójkąta. Natomiast odcinek, czy prosta nie są obszarami, gdyż nie są zbiorami otwartymi w tej przestrzeni. Zbiór

$$A=\{(x,y)\in R^2:\ 1<x^2<4\ \wedge\ -1<y<1\}$$

(przedstawiony na rysunku 1.4) także nie jest obszarem, gdyż jest wprawdzie zbiorem otwartym, ale nie jest zbiorem spójnym.



Rys 1.4

W przestrzeni R^2 koło wraz z okręgiem, trójkąt wraz z brzegiem są obszarami domkniętymi.

W przestrzeni R^3 obszarami są na przykład: wnętrze kuli, prostopadłościanu lub stożka. Czasem używa się w tym przypadku terminu: o b s z a r p r z e s t r z e n n y .

ZADANIA DO ROZWIĄZANIA.

1. Niech $X=R\times R$ i niech

   $$\rho(p_1,p_2)=|x_1-x_2|+|y_1-y_2|$$

   dla dowolnych $p_1=(x_1,y_1)$, $p_2=(x_2,y_2)\in X$. Wykazać, że funkcja ρ jest metryką i podać jej interpretację geometryczną.
2. Narysować kulę $K(0,1)$ w przestrzeni metrycznej z zadania 1.
3. Podać po dwa przykłady zbiorów ograniczonych i zbiorów nieograniczonych w przestrzeniach R^1, R^2 i R^3.

4. Podać po dwa przykłady zbiorów otwartych i zbiorów domkniętych w przestrzeniach R^1, R^2 i R^3.
5. Naszkicować zbiór A i wskazać punkty wewnętrzne, brzegowe i punkty skupienia tego zbioru, gdy:

   a) $A=\{x\in R:\ 0<x^2<4\}$, b) $A=\{x\in R:\ 0<x^2-1\leq 3\}$,

   c) $A=\{(x,y)\in R^2:\ |x-1|\leq 2\}$, d) $A=\{(x,y)\in R^2:\ y^2-1<3\}$,

   e) $A=\{(x,y)\in R^2:\ y>x^2\vee y=0\}$,

   f) $A=\{(x,y)\in R^2:\ 2x+3<x^2+y^2\leq 0\}$,

   g) $A=\{(x,y)\in R^2:\ y<3x-x^2\ \wedge\ y=x\}$,

   h) $A=\{(x,y)\in R^2:\ x^2-1\leq y<x^2\}$.
6. Czy zbiór A może być zbiorem otwartym, jeżeli wiadomo, że punkt $p_0\in A$ jest jego punktem brzegowym?
7. Podać przykład zbioru $A\subset R^2$, który nie ma punktów skupienia. Czy może to być zbiór otwarty?
8. Podać przykład zbioru $A\subset R^2$, którego każdy punkt jest punktem skupienia tego zbioru.
9. Podać przykłady podzbiorów przestrzeni R^1 i R^2, które nie są zbiorami otwartymi, ani też zbiorami domkniętymi.
10. Naszkicować zbiory A_k, $k=1,\ldots,7$, oraz określić wnętrze $\mathrm{Int}A_k$ i domknięcie $\overline{A}_k$ każdego z tych zbiorów, gdy

    $A_1=\{x\in R:\ 0\leq x<1\}$, $A_2=\{x\in R:\ -1\leq x\leq 2\}$,

    $A_3=\{(x,y)\in R^2:\ 1<x^2+1\leq 2\}$, $A_4=\{(x,y)\in R^2:\ 0<|y+1|<1\}$,

    $A_5=\{(x,y)\in R^2:\ 1\leq x\leq 2\ \wedge\ -\sqrt{x}<y\leq\sqrt{x}\}$,

    $A_6=\{(x,y)\in R^2:\ -2<x<1\ \wedge\ x^2+2x<y<x+2\}$,

    $A_7=\{(x,y)\in R^2:\ 2x-1<x^2+y^2\leq 2x\}$.
11. Podać przykłady zbiorów $A\subset R^2$ takich, że

    a) A jest obszarem,

    b) A jest zbiorem otwartym, ale nie jest obszarem,

    c) A jest zbiorem spójnym, ale nie jest obszarem.

12. Naszkicować zbiór A, a następnie odpowiedzieć na pytanie , czy zbiór ten jest obszarem (i dlaczego?):
a) $A=\{(x,y)\in R^2: x^2+y^2-2x<0 \vee x^2+y^2+4x<0\}$,
b) $A=\{(x,y)\in R^2: 4x-3<x^2+y^2<6x\}$,
c) $A=\{(x,y)\in R^2: y>x^2-1 \wedge y-x\le 0\}$.

13. Naszkicować zbiór $A\cap B$, gdy
$A=\{(x,y)\in R^2: 4<x^2+y^2<9\}$, $B=\{(x,y)\in R^2: 3x^2-y-3<0\}$.
Czy zbiór $A\cap B$ jest obszarem ?

14. Naszkicować zbiór $A\cap B$, gdy
$A=\{(x,y)\in R^2: 1\le x^2+y^2\le 4\}$, $B=\{(x,y)\in R^2: x-4y^2+1\ge 0\}$.
Czy zbiór $A\cap B$ jest a) zbiorem domkniętym, b) obszarem domkniętym ?

Odpowiedzi.

2. Wsk. $K(0,1)=\{p\in X:\rho(p,0)<1\}=\{(x,y)\in X:|x|+|y|<1\}$.

10. $\operatorname{Int}A_1=\{x\in R: 0<x<1\}$, $\overline{A}_1=\{x\in R: 0\le x\le 1\}$;
$\operatorname{Int}A_2=\{x\in R: -1<x<2\}$, $\overline{A}_2=A_2$;
$\operatorname{Int}A_3=\{(x,y)\in R^2: 1<x^2+1<2\}=\{(x,y)\in R^2: -1<x<0 \vee 0<x<1\}$,
$\overline{A}_3=\{(x,y)\in R^2: -1\le x\le 1\}$;
$\operatorname{Int}A_4=A_4=\{(x,y)\in R^2: -2<y<-1 \vee -1<y<0\}$,
$\overline{A}_4=\{(x,y)\in R^2: -2\le y\le 0\}$;
$\operatorname{Int}A_5=\{(x,y)\in R^2: 1<x<2 \wedge -\sqrt{x}<y<\sqrt{x}\}$,
$\overline{A}_5=\{(x,y)\in R^2: 1\le x\le 2 \wedge -\sqrt{x}\le y\le\sqrt{x}\}$;
$\operatorname{Int}A_6=A_6$, $\overline{A}_6=\{(x,y)\in R^2: -2\le x\le 1 \wedge x^2+2x\le y\le x+2\}$;
$\operatorname{Int}A_7=\{(x,y)\in R^2: 2x-1<x^2+y^2<2x\}=\{(x,y)\in R^2: 0<(x-1)^2+y^2<1\}$,
$\overline{A}_7=\{(x,y)\in R^2: 2x-1\le x^2+y^2\le 2x\}=\{(x,y)\in R^2: (x-1)^2+y^2\le 1\}$.
Sporządzenie rysunków pozostawiamy Czytelnikowi.

12. a) nie, b) tak, c) nie. 13. nie. 14. a) tak, b) nie.

2. ZBIEŻNOŚĆ CIĄGU W PRZESTRZENI METRYCZNEJ.

GRANICA CIĄGU. Niech (X,ρ) będzie dowolną przestrzenią metryczną, a (p_n) dowolnym ciągiem elementów zbioru X. Niech p_0 również należy do X.

Mówimy, że **ciąg (p_n) ma granicę p_0** (lub: **ciąg (p_n) jest zbieżny do p_0**), co zapisujemy

$$\lim_{n\to\infty} p_n = p_0 \qquad \text{lub} \qquad p_n \to p_0 \ \text{ przy } \ n\to\infty,$$

gdy ciąg $(\rho(p_n,p_0))$ ma granicę równą zeru, czyli

$$\lim_{n\to\infty} p_n = p_0 \overset{\text{def}}{\Leftrightarrow} \lim_{n\to\infty} \rho(p_n,p_0)=0. \tag{2.1}$$

Ponieważ

$$\lim_{n\to\infty} \rho(p_n,p_0)=0 \Leftrightarrow \bigwedge_{\varepsilon>0}\bigvee_{K}\bigwedge_{n>K} \rho(p_n,p_0)<\varepsilon,$$

więc

$$\lim_{n\to\infty} p_n = p_0 \Leftrightarrow \bigwedge_{\varepsilon>0}\bigvee_{K}\bigwedge_{n>K} \rho(p_n,p_0)<\varepsilon. \tag{2.2}$$

Ciąg, który nie jest zbieżny nazywamy **rozbieżnym**.

Na przykład dla ciągu (p_n), gdy $p_n=(x_n,y_n)\in R^2$, $n\in N$, i $p_0=(x_0,y_0)\in R^2$ otrzymujemy

$$\lim_{n\to\infty} p_n = p_0 \Leftrightarrow \bigwedge_{\varepsilon>0}\bigvee_{K}\bigwedge_{n>K} \sqrt{(x_n-x_0)^2+(y_n-y_0)^2}<\varepsilon,$$

gdyż

$$\rho(p_n,p_0)=\sqrt{(x_n-x_0)^2+(y_n-y_0)^2}.$$

PRZYKŁAD 2.1. Niech (p_n) będzie ciągiem elementów przestrzeni R^2, przy czym

$$p_n=\left(\frac{n+3}{n},\frac{4}{n}\right), \quad n\in N.$$

Wykażemy, że granicą tego ciągu (p_n) jest $p_0=(1,0)$.

Istotnie, ponieważ

$$\rho(p_n,p_0)=\sqrt{(\frac{n+3}{n}-1)^2+(\frac{4}{n}-0)^2}=\frac{5}{n},\quad n\in N,$$

więc

$$\lim_{n\to\infty}\rho(p_n,p_0)=\lim_{n\to\infty}\frac{5}{n}=0.$$

Oznacza to, zgodnie z definicją (2.1), że $p_n\to p_0$. ■

Dla ciągów, których wyrazy są elementami dowolnej przestrzeni metrycznej pozostają prawdziwe następujące twierdzenia, poznane już wcześniej dla ciągów o wyrazach rzeczywistych:

TWIERDZENIE 2.1. Każdy podciąg ciągu zbieżnego jest zbieżny i ma tę samą granicę, co dany ciąg.

TWIERDZENIE 2.2. Ciągi różniące się skończoną liczbą wyrazów są jednocześnie zbieżne (i mają wtedy tę samą granicę) lub jednocześnie rozbieżne.

TWIERDZENIE 2.3. Każdy ciąg zbieżny jest ograniczony.

D o w ó d . Załóżmy, że ciąg (p_n) jest zbieżny i ma granicę p_0. Wówczas $\rho(p_n,p_0)\to 0$, czyli

$$\bigwedge_{\varepsilon>0}\bigvee_{K}\bigwedge_{n>K}\rho(p_n,p_0)<\varepsilon.$$

Zatem dla $\varepsilon=1$ istnieje taka liczba $K\in N$, że dla $n>K$ zachodzi nierówność $\rho(p_n,p_0)<1$.

Niech

$$r_0=\max\{1,\rho(p_1,p_0),\rho(p_2,p_0),\ldots,\rho(p_K,p_0)\}.$$

Stąd wynika, że $\rho(p_n,p_0)\le r_0$ dla $n\in N$, a w konsekwencji dla każdego $n\in N$ i $r>r_0$ zachodzi warunek $\rho(p_n,p_0)<r$, czyli

$$p_n\in K(p_0,r).$$

Oznacza to, że ciąg (p_n) jest ograniczony. □

ZBIEŻNOŚĆ CIĄGU W PRZESTRZENI R^k. Na początek przyjmijmy, że (p_n) jest ciągiem elementów przestrzeni R^2. Zatem $p_n=(x_n,y_n)$, gdzie $x_n\in R$, $y_n\in R$, $n\in N$. Na przykład ciągi (p_n) i (q_n), gdy

$$p_n=(1+3^{-n},\sqrt[n]{n}-2),\quad q_n=(\frac{n^2}{3-n^2},\frac{n^2+1}{3+n}),\quad n\in N,$$

są takimi ciągami.

Podamy teraz twierdzenie, które pozwala nam problem obliczania granic takich ciągów sprowadzić do obliczania granic ciągów o wyrazach rzeczywistych.

TWIERDZENIE 2.4. Załóżmy, że $p_n=(x_n,y_n)\in R^2$, $n\in N$, i $p_0=(x_0,y_0)\in R^2$. Wówczas

$$\lim_{n\to\infty}p_n=p_0\iff(\lim_{n\to\infty}x_n=x_0\wedge\lim_{n\to\infty}y_n=y_0).$$

D o w ó d $\Rightarrow$. Niech ε będzie dowolnie ustaloną liczbą dodatnią. Ponieważ $p_n\to p_0$, więc

$$\bigvee_{K}\bigwedge_{n>K}\rho(p_n,p_0)<\varepsilon,$$

czyli

$$\sqrt{(x_n-x_0)^2+(y_n-y_0)^2}<\varepsilon\quad\text{dla}\quad n>K.$$

Stąd wynika, że

$$|x_n-x_0|<\varepsilon\quad\text{i}\quad|y_n-y_0|<\varepsilon\quad\text{dla}\quad n>K.$$

Zatem

$$\bigwedge_{\varepsilon>0}\bigvee_{K}\bigwedge_{n>K}|x_n-x_0|<\varepsilon\quad\text{oraz}\quad\bigwedge_{\varepsilon>0}\bigvee_{K}\bigwedge_{n>K}|y_n-y_0|<\varepsilon,$$

a to oznacza, że $x_n\to x_0$ oraz $y_n\to y_0$.

D o w ó d $\Leftarrow$. Niech ε będzie dowolnie ustaloną liczbą dodatnią. Ponieważ $x_n\to x_0$ i $y_n\to y_0$, więc

$$\bigvee_{K_1}\bigwedge_{n>K_1}|x_n-x_0|<\frac{\varepsilon}{\sqrt{2}}\quad\text{oraz}\quad\bigvee_{K_2}\bigwedge_{n>K_2}|y_n-y_0|<\frac{\varepsilon}{\sqrt{2}}.$$

W konsekwencji, dla $n > K = \max\{K_1, K_2\}$, mamy

$$\rho(p_n, p_0) = \sqrt{(x_n - x_0)^2 + (y_n - y_0)^2} < \sqrt{(\frac{\varepsilon}{\sqrt{2}})^2 + (\frac{\varepsilon}{\sqrt{2}})^2}$$

Wykazaliśmy więc, że

$$\bigwedge_{\varepsilon>0} \bigvee_{K} \bigwedge_{n>K} \rho(p_n, p_0) < \varepsilon,$$

a to oznacza, że $p_n \to p_0$. □

Ogólniej:

TWIERDZENIE 2.4'. Załóżmy, że $p_n = (x_{1n}, \ldots, x_{kn}) \in R^k$, $n \in N$, i $p_0 = (x_{10}, \ldots, x_{k0}) \in R^k$, $k \geq 2$. Wówczas

$$\lim_{n\to\infty} p_n = p_0 \Leftrightarrow \bigwedge_{i=1,\ldots,k} (\lim_{n\to\infty} x_{in} = x_{i0}).$$

U w a g a. Z twierdzenia tego wynika, że ciąg (p_n) jest zbieżny wtedy i tylko wtedy, gdy każdy z ciągów (x_{in}), $i = 1, \ldots, k$, jest zbieżny. Tak więc, jeżeli przynajmniej jeden z ciągów (x_{in}) jest rozbieżny, to ciąg (p_n) jest rozbieżny.

P R Z Y K Ł A D 2.2. Obliczamy granice:

a) $\lim_{n\to\infty}(1 - 3^{-n}, \sqrt[n]{n} - 2) = (1, -1)$, gdyż $\lim_{n\to\infty}(1 - 3^{-n}) = 1$ i $\lim_{n\to\infty}(\sqrt[n]{n} - 2) = -1$;

b) $\lim_{n\to\infty}(\frac{2}{n+1}, \frac{1+2^n}{e^n}, \sqrt[n]{3 - n + n^3}) = (0,0,1)$, gdyż $\lim_{n\to\infty}\frac{2}{n+1} = 0$, $\lim_{n\to\infty}\frac{1+2^n}{e^n} = \lim_{n\to\infty}(\frac{1}{e^n} + (\frac{2}{e})^n) = 0$, $\lim_{n\to\infty}\sqrt[n]{3 - n + n^3} = 1$;

c) ciąg $(3^{-n}, \cos n, \sqrt[n]{n})$ jest rozbieżny, gdyż ciąg $(\cos n)$ jest rozbieżny. ■

ZADANIA DO ROZWIĄZANIA.

1. Obliczyć granicę ciągu (p_n), gdy:

   a) $p_n = (\frac{3+n^2}{2+n^2}, \frac{-2n}{1+n})$, b) $p_n = (\frac{1}{2^n}, \frac{1-3^{n+1}}{2^n+3^n})$,

   c) $p_n = (2 + 3^{1-n}, \sqrt[n]{5^{n+1} - 2^n}, \sqrt[n]{2 - n^2 + 3n^4} - 3)$,

   d) $p_n = ((\frac{n+1}{n})^{2n}, (\frac{2n}{1+n})^{-n}, (\frac{3}{n})^n)$.

2. Podać przykład ciągu (p_n) takiego, że

   a) $p_n \in R^2, n \in N$ i $p_n \to (-1,2)$, b) $p_n \in R^2, n \in N$ i $p_n \to (0,1)$,

   c) $p_n \in R^3, n \in N$ i $p_n \to (0,1,0)$, d) $p_n \in R^3, n \in N$ i $p_n \to (-1,1,2)$.

3. Podać przykład ciągu rozbieżnego (p_n) takiego, że

   a) $p_n \in R^2$, $n \in N$, b) $p_n \in R^3$, $n \in N$.

O d p o w i e d z i.

1. a) $(1,-2)$, b) $(0,-3)$, c) $(2,5,-2)$, d) $(e^2,0,0)$.

3. FUNKCJE WIELU ZMIENNYCH.

FUNKCJE WIELU ZMIENNYCH RZECZYWISTYCH. W pierwszym tomie tej książki (rozdz. I, 3) podane zostały podstawowe wiadomości o funkcjach. Przypomnijmy krótko: funkcja odwzorowująca zbiór X w zbiór Y (odwzorowanie zbioru X w zbiór Y) jest to przyporządkowanie każdemu elementowi zbioru X dokładnie jednego elementu zbioru Y. Funkcję f odwzorowującą zbiór X w zbiór Y zapisujemy

$$f: X \to Y.$$

Funkcje postaci $f: D \to R$, $D \subset R$, są funkcjami o wartościach rzeczywistych jednej zmiennej rzeczywistej i były już przedmiotem naszych wcześniejszych rozważań.

Funkcje postaci $f: D \to R$, $D \subset R^n$ $(n \geq 2)$ są funkcjami określonymi na pewnych podzbiorach przestrzeni R^n, zatem ich argumentami są punkty $p = (x_1, x_2, \ldots, x_n) \in R^n$. Funkcje takie nazywamy **funkcjami rzeczywistymi n zmiennych rzeczywistych** i zapisujemy: $f(x_1, x_2, \ldots, x_n)$, gdzie $(x_1, x_2, \ldots, x_n) \in D$ lub krócej: $f(p)$, gdzie $p = (x_1, x_2, \ldots, x_n) \in D$.

PRZYKŁAD 3.1.

a) Funkcje określone wzorami

$$f(x,y) = \sqrt{1 - x^2 - y^2}, \qquad g(x,y,z) = z^2 - \ln(1 + x^2 + y^2),$$

gdzie x, y, z oznaczają zmienne rzeczywiste, są funkcjami o wartościach rzeczywistych odpowiednio dwóch i trzech zmiennych rzeczywistych, przy czym dziedziną funkcji f jest koło domknięte

$$D_1 = \{(x,y) \in R^2 : x^2 + y^2 \leq 1\},$$

a dziedziną funkcji g jest zbiór $D_2 = R^3$.

b) Funkcja określona wzorem

$$\varphi(x,y) = (x^2 - y^2,\ xy,\ \sqrt{5 + y^2}), \quad (x,y) \in R^2,$$

jest funkcją dwóch zmiennych rzeczywistych o wartościach z przestrzeni R^3. Dla przykładu: $\varphi(-1,1) = (0,-1,\sqrt{6})$, $\varphi(0,2) = (-4,0,3)$.

Funkcje, których wartości należą do zbioru R^n, $n > 1$, nazywane są **funkcjami wektorowymi**. ■

Obecnie zajmować się będziemy funkcjami o wartościach rzeczywistych n zmiennych rzeczywistych (nazywanymi krócej funkcjami wielu zmiennych), ze szczególnym uwzględnieniem funkcji dwóch zmiennych. Zatem przedmiotem naszych rozważań będą funkcje $f: D \to R$, $D \subset R^n$, gdy $n \geq 2$.

Funkcję określoną wzorem będziemy zawsze rozważać w dziedzinie naturalnej, chyba, że dziedzina będzie wyraźnie inaczej określona. Przypomnijmy, że dziedzina naturalna funkcji określonej wzorem jest to zbiór tych argumentów, dla których wzór ten ma sens.

PRZYKŁAD 3.2. Wyznaczymy dziedzinę D funkcji $z = z(x,y)$ określonej wzorem:

a) $z = \sqrt{x} + \sqrt{2 - y}$,

b) $z = \ln(2 - x) - y\ln(x + y)$,

c) $z = 3\arccos(2y - 3)$,

d) $z = \ln(2y - y^2) + \ln(y - \ln x)$,

e) $z = \dfrac{1 + \sqrt{4x - x^2 - y}}{\sqrt{y + 2x - x^2}}$,

f) $z = \dfrac{y\ln(2 + x - x^2)}{\sqrt{9 - x^2 - y^2}}$,

g) $z = \dfrac{\sqrt{2x - x^2 - y^2}}{2 + \sqrt{y + x}}$,

h) $z = \dfrac{\ln(4y - x^2 - y^2)}{x\sqrt{x^2 + y^2 - 2y}}$,

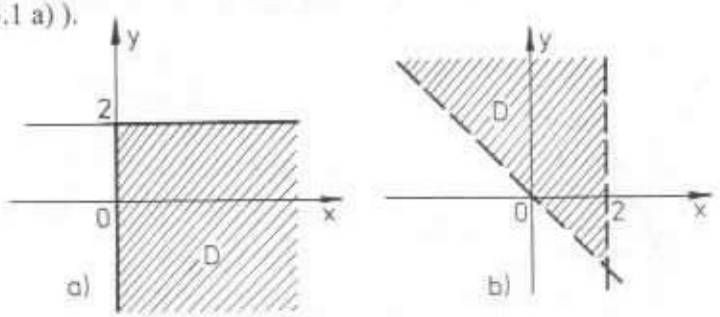
i) $z = x\sqrt{-x^2 - y^2}$,

j) $z = y\sqrt{y - x} - 2\sqrt{x - y}$.

a) Dziedziną tej funkcji jest zbiór $D \subset R^2$ punktów (x,y), dla których $x \geq 0$ i $2 - y \geq 0$, czyli

$$D = \{(x,y) \in R^2 : x \geq 0 \wedge y \leq 2\},$$

(rys. 3.1 a)).



Rys 3.1

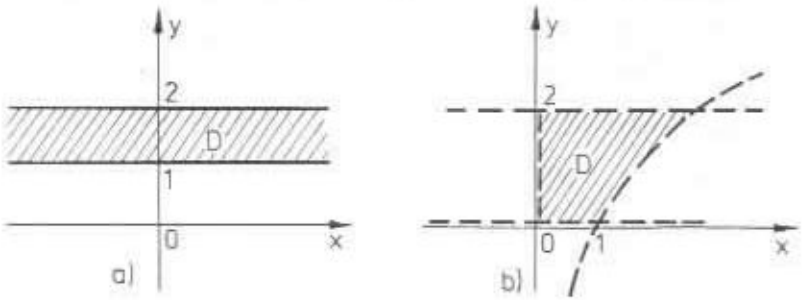
b) $D = \{(x,y) \in R^2 : 2 - x > 0 \wedge x + y > 0\} =$

$$= \{(x,y) \in R^2 : x < 2 \wedge y > -x\}.$$

Zbiór D przedstawiony jest na rysunku 3.1 b).

c) Dziedziną funkcji jest zbiór $D \subset R^2$ punktów (x,y), dla których $-1 \leq 2y - 3 \leq 1$, czyli $1 \leq y \leq 2$. Zatem (rys. 3.2 a))

$$D = \{(x,y) \in R^2 : 1 \leq y \leq 2\},$$



Rys 3.2

d) $D=\{(x,y)\in R^2:\ 2y-y^2>0 \wedge y-\ln x>0 \wedge x>0\}=$

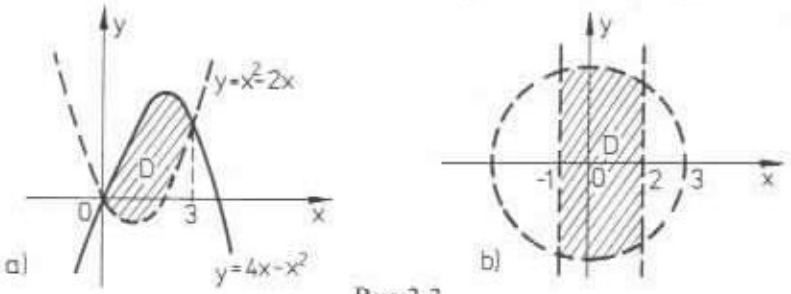
$=\{(x,y)\in R^2:\ 0<y<2 \wedge x>0 \wedge y>\ln x\}$,

(rys. 3.2 b)).

e) Dziedziną tej funkcji jest zbiór $D\subset R^2$ punktów (x,y), dla których $y+2x-x^2>0$ i $4x-x^2-y\geq 0$, czyli (rys. 3.3 a))

$D=\{(x,y)\in R^2:\ y>x^2-2x \wedge y\leq 4x-x^2\}=$

$=\{(x,y)\in R^2:\ x^2-2x<y\leq 4x-x^2\}$.



Rys 3.3

f) $D=\{(x,y)\in R^2:\ 2+x-x^2>0 \wedge 9-x^2-y^2>0\}=$

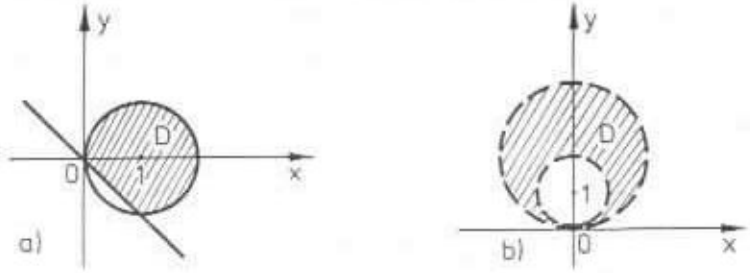
$=\{(x,y)\in R^2:\ -1<x<2 \wedge x^2+y^2<9\}$.

Obszar D pokazany jest na rysunku 3.3 b).

g) $D=\{(x,y)\in R^2:\ 2x-x^2-y^2\geq 0 \wedge y+x\geq 0\}=$

$=\{(x,y)\in R^2:\ (x-1)^2+y^2\leq 1 \wedge y\geq -x\}$,

(rys. 3.4 a)).



Rys 3.4

h) $D=\{(x,y)\in R^2:\ 4y-x^2-y^2>0 \wedge x^2+y^2-2y>0\}=$

$=\{(x,y)\in R^2:\ x^2+(y-2)^2<4 \wedge x^2+(y-1)^2>1\}$,

(rys. 3.4 b)).

i) $D=\{(x,y)\in R^2:\ -x^2-y^2\geq 0\}=\{(x,y)\in R^2:\ x^2+y^2=0\}=$

$=\{(0,0)\}$.

j) $D=\{(x,y)\in R^2:\ y-x\geq 0 \wedge x-y\geq 0\}=\{(x,y)\in R^2:\ y=x\}$.

Interpretację geometryczną zbiorów D w punktach i) , j) pozostawiamy Czytelnikowi. ■

WYKRESY FUNKCJI DWÓCH ZMIENNYCH. Wykres funkcji dwóch zmiennych określonej wzorem $z=f(x,y)$, $(x,y)\in D$ jest to zbiór $W\subset R^3$ określono następująco

$W=\{(x,y,z)\in R^3:\ (x,y)\in D \wedge z=f(x,y)\}$.

Wykresem funkcji dwóch zmiennych jest więc pewien podzbiór przestrzeni R^3. W przypadku gdy funkcja jest określona na pewnym obszarze i jest ciągła na tym obszarze, wykresem jest pewna powierzchnia. W ogólności tak być nie musi (por. przykład 3.3. e) i f)).

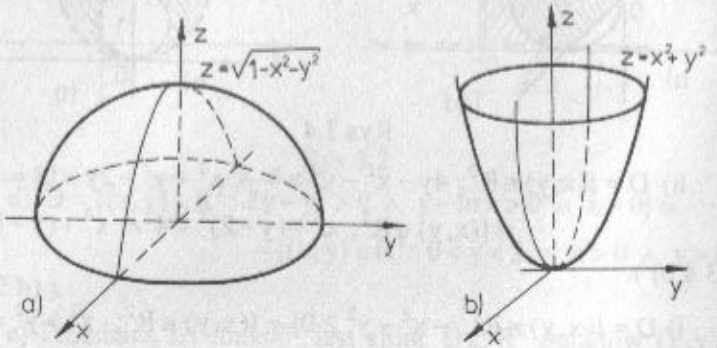
PRZYKŁAD 3.3.

a) Funkcja

$$z=\sqrt{1-x^2-y^2}$$

jest określona na kole $D=\{(x,y)\in R^2:\ x^2+y^2\leq 1\}$, a jej wykresem jest "górna połowa" powierzchni kulistej o środku w punkcie $(0,0,0)$ i promieniu $r=1$ (rys 3.5 a)).

Równanie $x^2+y^2+z^2=1$ określa powierzchnię kulistą (sferę) o środku w punkcie (0,0,0) i promieniu r = 1. Z równania tego otrzymujemy $z=\sqrt{1-x^2-y^2}$, gdy $z\geq 0$ lub $z=-\sqrt{1-x^2-y^2}$, gdy $z\leq 0$, przy czym pierwsze z tych równań określa "górną połowę", a drugie równanie - "dolną połowę" powierzchni kulistej.



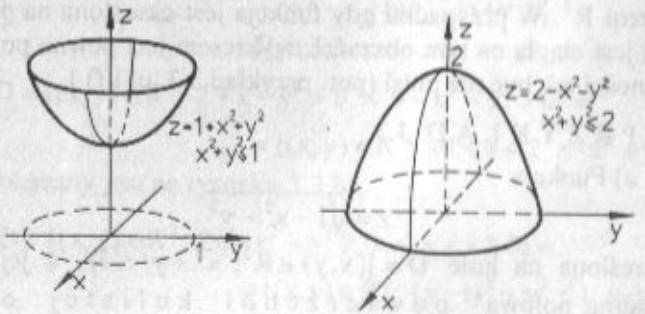
Rys 3.5

b) Funkcja

$$z=x^2+y^2$$

jest określona dla $(x,y)\in R^2$, a jej wykresem jest powierzchnia nazywana paraboloidą obrotową (rys 3.5 b)).

Powierzchnię tę można otrzymać przez obrót dokoła osi 0z paraboli o równaniu $z=x^2$ (leżącej w płaszczyźnie 0xz). Przekrój tej powierzchni dowolną płaszczyzną z=a, a > 0, jest okręgiem, a przekroje płaszczyznami x = b lub y = b są parabolami. Ogólniej: równanie $z=\alpha x^2+\beta y^2$, $\alpha>0$, $\beta>0$ określa paraboloidę eliptyczną, której przekroje dowolną płaszczyzną z = a, a > 0, są elipsami (por. rozdz. I, 5).



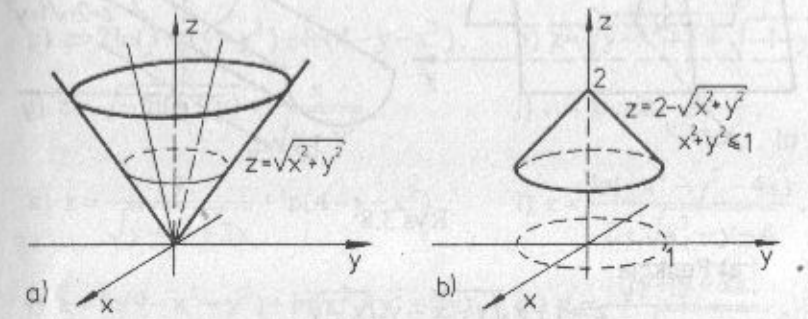
Rys 3.6

Na rysunku 3.6. przedstawione są wykresy funkcji $z=1+x^2+y^2$ dla $x^2+y^2\leq 1$ oraz $z=2-x^2-y^2$ dla $x^2+y^2\leq 2$.

c) Funkcja

$$z=\sqrt{x^2+y^2}$$

jest określona dla $(x,y)\in R^2$. Wykresem tej funkcji (rys 3.7a)) jest "górna połowa" stożka $z^2=x^2+y^2$ (dla $z\geq 0$).



Rys 3.7

Na rysunku 3.7 b) przedstawiony jest wykres funkcji $z=2-\sqrt{x^2+y^2}$ dla $x^2+y^2\leq 1$.

Stożek $z^2=x^2+y^2$ można otrzymać przez obrót dokoła osi 0z prostej z = x zawartej w płaszczyźnie 0xz. Przekrój tej powierzchni dowolną płaszczyzną z = a, $a\neq 0$ jest okręgiem. Ogólniej: równanie $z^2=\alpha x^2+\beta y^2$, $\alpha>0$, $\beta>0$, określa stożek, którego przekroje dowolną płaszczyzną z = a, $a\neq 0$ są elipsami (por. rozdz. I, 5).

d) Wykresem funkcji z = z(x,y) określonej wzorem

$$z=1-x^2,\quad (x,y)\in R^2,$$

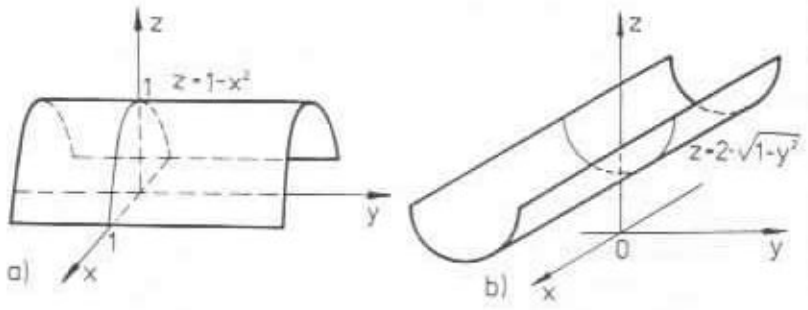
jest powierzchnia walcowa, której kierownicą jest parabola $z=1-x^2$ zawarta w płaszczyźnie 0xz, a tworzące są równoległe do osi 0y (rys 3.8 a)).

Na rysunku 3.8 b) przedstawiony jest wykres funkcji z = z(x,y) określonej wzorem

$$z=2-\sqrt{1-y^2}$$

dla $(x,y)\in D=\{(x,y)\in R^2:\ -1\leq y\leq 1\}$. Wykresem tej funkcji jest

"dolna połowa" powierzchni walcowej, której kierownicą jest okrąg o równaniu $y^2+(z-2)^2=1$ zawarty w płaszczyźnie 0yz, a tworzące są równoległe do osi 0x.



Rys 3.8

e) Funkcja

$$z=1+x\sqrt{y-x}-y\sqrt{x-y}$$

jest określona na zbiorze $D=\{(x,y)\in R^2:\ y-x\geq 0 \wedge x-y\leq 0\}=$ $=\{(x,y)\in R^2:\ y=x\}$. Oznacza to, że funkcja ta jest określona jedynie w tych punktach płaszczyzny 0xy, które leżą na prostej $y=x$, przy czym w każdym takim punkcie wartość funkcji jest równa 1.

Wykresem tej funkcji jest prosta o równaniach: $x=t$, $y=t$, $z=1$, $t\in R$. Sporządzenie rysunku pozostawiamy Czytelnikowi.

f) Funkcja

$$z=2+\sqrt{-x^2-y^2}$$

jest określona jedynie w punkcie $(x,y)=(0,0)$. W konsekwencji wykres tej funkcji składa się z jednego punktu. Jest to punkt $P(0,0,2)$. ■

ZADANIA DO ROZWIĄZANIA.

1. Podać przykład funkcji $f: X\to Y$, gdy:

a) $X=R$, $Y=R$, b) $X=R$, $Y=R^2$, c) $X=R^2$, $Y=R$,

d) $X=R^2$, $Y=R^2$, e) $X=R^3$, $Y=R$, f) $X=R^3$, $Y=R^2$.

2. Wyznaczyć dziedzinę D (narysować ten zbiór i zapisać nierównością mi) funkcji $z=z(x,y)$, gdy:

a) $z=\dfrac{y^2-1}{\sqrt{1-x}}+\ln(x-y)$, b) $z=\dfrac{\ln(3+y)}{\sqrt{2x-y-x^2}}$,

c) $z=\ln(4x-x^2-y^2)+x\ln(|x+1|-2)$, d) $z=3x^2-\sqrt{4-|1-y|}$,

e) $z=\arcsin(2x-3)+\sqrt{9-x^2-y^2}$, f) $z=\sqrt{x}-3\arcsin\dfrac{y}{x}$,

g) $z=2\ln(3+2y-y^2)+\ln(4-y-x^2)$, h) $z=\sqrt{y-x^2+1}+\sqrt{-1-y}$,

i) $z=\sqrt{-\ln(x+y)}+\dfrac{x}{\sqrt{y-y^2}}$, j) $z=\arcsin\dfrac{1}{x^2+y^2}$,

k) $z=\dfrac{1}{\sqrt{y-x^2+2x}}+\ln(4-y-x^2)$, l) $z=\dfrac{\ln(-x^2-y^2-4x)}{2+\sqrt{x^2-y-4}}$,

ł) $z=\ln(9-x^2-y^2)+\ln(x^2+y^2-2y)$, m) $z=\dfrac{\sqrt{y-|y-2x|}}{1+\sqrt{2-y-x^2}}$,

n) $z=x\ln(9-y^2)+\sqrt{3-|y+1|}$, o) $z=y^3\arccos(1-|x^2-x|)$,

p) $z=\sqrt{y-1}+3\sqrt{\cos x-y}$, r) $z=2y\arccos(1+x^2)$.

3. Podać przykład funkcji $z=z(x,y)$, której dziedziną naturalną jest zbiór D:

a) $D=\{(x,y)\in R^2:\ y<x^2+1\}$, b) $D=\{(x,y)\in R^2:\ y=x^2\}$,

c) $D=\{(x,y)\in R^2:\ x^2+y^2-2x\leq 0\}$, d) $D=\{(1,0)\}$.

4. Naszkicować wykres funkcji $z=z(x,y)$, gdy:

a) $z=4-4y-2x$, b) $z=2$, gdy $x^2+y^2\leq 1$,

c) $z=-x^2-y^2$, d) $z=1-x^2-y^2$,

e) $4z=x^2+4y^2$, f) $z=1+\sqrt{x^2+y^2}$,

g) $z=\sqrt{x^2+y^2}-2$, h) $z=4-x^2$,

i) $z=y^2-y$, j) $z=\sqrt{4-x^2}$,

k) $z=2+\sqrt{4-x^2}$, l) $z=1+\sqrt{y-x^2+1}+\sqrt{-1-y}$,

ł) $z=2-\sqrt{y-x^2+x}+y^2\sqrt{x^2-y-x}$,

m) $z=x-\sqrt{y-x^2}-y\sqrt{x^2-y}+3\sqrt{2y-y^2-1}+3y^2$.

Odpowiedzi.

2. a) $D=\{(x,y)\in R^2: x<1 \wedge y<x\}$, b) $D=\{(x,y)\in R^2: -3<y<2x-x^2\}$,

c) $D=\{(x,y)\in R^2: (x-2)^2+y^2<4 \wedge x>1\}$,

d) $D=\{(x,y)\in R^2: -3\le y\le 5\}$, e) $D=\{(x,y)\in R^2: 1\le x\le 2 \wedge x^2+y^2\le 9\}$,

f) $D=\{(x,y)\in R^2: -x\le y\le x \wedge x>0\}$,

g) $D=\{(x,y)\in R^2: y<4-x^2 \wedge -1<y<3\}=$

$=\{(x,y)\in R^2: -1<y<3 \wedge -\sqrt{4-y}<x<\sqrt{4-y}\}$, h) $D=\{(0,-1)\}$,

i) $D=\{(x,y)\in R^2: -y<x\le 1-y \wedge 0<y<1\}$, j) $D=\{(x,y)\in R^2: x^2+y^2\ge 1\}$,

k) $D=\{(x,y)\in R^2: x^2-2x<y<4-x^2\}$,

l) $D=\{(x,y)\in R^2: (x+2)^2+y^2<4 \wedge y\le x^2-4\}$,

ł) $D=\{(x,y)\in R^2: x^2+y^2<9 \wedge x^2+(y-1)^2>1\}$,

m) $D=\{(x,y)\in R^2: 0\le x\le 1 \wedge x\le y\le 2-x^2\}$, n) $D=\{(x,y)\in R^2: -3<y\le 2\}$,

o) $D=\{(x,y)\in R^2: -1\le x\le 2\}$,

p) $D=\{(x,y)\in R^2: y\le\cos x \wedge y\ge 1\}=\{(2k\pi,1): k\in C\}$, r) $D=\{(x,y)\in R^2: x=0\}$.

Sporządzenie odpowiednich rysunków pozostawiamy Czytelnikowi.

4. Wsk. a) płaszczyzna przechodząca przez punkty (2,0,0), (0,1,0), (0,0,4);

b) część płaszczyzny $z=2$, której rzutem na pł. 0xy jest koło: $x^2+y^2\le 1$;

c) i d) paraboloidy obrotowe; e) paraboloida eliptyczna; f) i g) stożki;

h) powierzchnia walcowa, której kierownicą jest parabola $z=4-x^2$ zawarta w pł. 0xz, a tworzące są równoległe do osi 0y; i) powierzchnia walcowa, której kierownicą jest parabola $z=y^2-y$ zawarta w pł. 0yz, a tworzące są równoległe do osi 0x; j) powierzchnia walcowa, której kierownicą jest półokrąg $x^2+z^2=4$, $z\ge 0$ zawarty w pł. 0xz, a tworzące są równoległe do osi 0y;

k) powierzchnia z punktu j) przesunięta o wektor [0,0,2]; l) funkcja jest określona na zbiorze jednoelementowym $D=\{(0,-1)\}$, przy czym $z(0,-1)=1$; wykresem jest zbiór jednoelementowy $W=\{(0,-1,1)\}$; ł) dziedziną funkcji jest zbiór $D=\{(x,y)\in R^2: y=x^2-x\}$, przy czym w każdym punkcie dziedziny $z(x,y)=2$; m) $D=\{(-1,1),(1,1)\}$, przy czym $z(-1,1)=2$, $z(1,1)=4$; wykres jest zbiorem dwuelementowym $W=\{(-1,1,2),(1,1,4)\}$.

4. GRANICA I CIĄGŁOŚĆ FUNKCJI WIELU ZMIENNYCH.

GRANICA FUNKCJI. Załóżmy, że D jest podzbiorem pewnej przestrzeni metrycznej X, a p_0 jest punktem skupienia tego zbioru (punkt p_0 może do zbioru należeć lub nie). Łatwo wykazać, że istnieje wówczas ciąg (p_n) spełniający warunki:

(4.1) $\quad p_n\in D$ i $p_n\ne p_0$ dla $n\in N$ oraz $p_n\to p_0$.

Istotnie, zgodnie z definicją punktu skupienia zbioru, w każdym otoczeniu punktu p_0 znajdują się elementy zbioru D, różne od p_0. Zatem dla każdego $n\in N$ istnieje punkt $p_n\in D$, $p_n\ne p_0$ taki, że $\rho(p_n,p_0)<\frac{1}{n}$. Stąd wynika, że $\rho(p_n,p_0)\to 0$, a to oznacza, że $p_n\to p_0$.

Niech X i Y będą dowolnymi przestrzeniami metrycznymi. Załóżmy, że $f: D\to Y$, przy czym $D\subset X$. Załóżmy, że p_0 jest punktem skupienia zbioru D oraz $g\in Y$.

Niżej podamy definicje Heinego i Cauchy'ego granicy g funkcji f w punkcie p_0. Wykazuje się, że definicje te są równoważne.

DEFINICJA HEINEGO. Mówimy, że funkcja **f ma w punkcie p_0 granicę g**, gdy dla dowolnego ciągu (p_n) takiego, że $p_n\in D$ i $p_n\ne p_0$ dla $n\in N$ i $p_n\to p_0$, ciąg $(f(p_n))$ ma granicę równą g, czyli

$$\lim_{p\to p_0} f(p)=g \overset{\text{def}}{\Leftrightarrow} \bigwedge_{(p_n)}((p_n\in D-\{p_0\},\ n\in N \wedge p_n\to p_0)\Rightarrow(f(p_n)\to g)).$$

DEFINICJA CAUCHY'EGO. Mówimy, że funkcja **f ma w punkcie p_0 granicę g**, gdy dla dowolnej liczby $\varepsilon>0$ istnieje taka liczba $\delta>0$, że dla punktów $p\in D$ spełniających nierówność $0<\rho(p,p_0)<\delta$, wartości funkcji spełniają nierówność $\rho(f(p),g)<\varepsilon$, czyli

$$\lim_{p\to p_0} f(p)=g \overset{\text{def}}{\Leftrightarrow} \bigwedge_{\varepsilon>0}\bigvee_{\delta>0}\bigwedge_{p\in D}(0<\rho(p,p_0)<\delta \Rightarrow \rho(f(p),g)<\varepsilon).$$

W szczególności, gdy funkcja f przyjmuje wartości rzeczywiste w definicji Cauchy'ego można uwzględnić, że

$$\rho(f(p),g)=|f(p)-g|,$$

a w przypadku gdy jest to funkcja n zmiennych rzeczywistych i $p=(x_1,x_2,\ldots,x_n)$, $p_0=(x_1^0,x_2^0,\ldots,x_n^0)$, mamy

$$\rho(p,p_0)=\sqrt{\sum_{i=1}^{n}(x_i-x_i^0)^2}.$$

Na przykład dla funkcji o wartościach rzeczywistych dwóch zmiennych rzeczywistych $z=f(x,y)$, $(x,y)\in D\subset R^2$, definicję Cauchy'ego granicy w punkcie $p_0=(x_0,y_0)$ można zapisać jak następuje:

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)=g \overset{\text{def}}{\Leftrightarrow}$$

$$\Leftrightarrow \bigwedge_{\varepsilon>0}\bigvee_{\delta>0}\bigwedge_{(x,y)\in D}(0<\sqrt{(x-x_0)^2+(y-y_0)^2}<\delta \Rightarrow |f(x,y)-g|<\varepsilon).$$

Dla funkcji o wartościach rzeczywistych można mówić także o granicy niewłaściwej $+\infty$ lub $-\infty$. Na przykład:

DEFINICJA HEINEGO.

$$\lim_{p\to p_0} f(p)=+\infty \overset{\text{def}}{\Leftrightarrow} \bigwedge_{(p_n)}((p_n\in D-\{p_0\}, n\in N \wedge p_n\to p_0)\Rightarrow(f(p_n)\to+\infty)).$$

DEFINICJA CAUCHY'EGO.

$$\lim_{p\to p_0} f(p)=+\infty \overset{\text{def}}{\Leftrightarrow} \bigwedge_{M}\bigvee_{\delta>0}\bigwedge_{p\in D}(0<\rho(p,p_0)<\delta \Rightarrow f(p)>M).$$

U w a g a .Definiując granicę funkcji jednej zmiennej w punkcie p_0 zakładaliśmy, że funkcja jest określona na pewnym sąsiedztwie tego punktu, a obecnie założyliśmy jedynie, że p_0 jest punktem skupienia dziedziny rozpatrywanej funkcji. To "osłabienie" założenia pozwala nam na przykład rozważać granicę funkcji określonej na obszarze D w punktach brzegowych tego obszaru mimo, że w żadnym sąsiedztwie takiego punktu funkcja nie musi być określona.

PRZYKŁAD 4.1. Weźmy pod uwagę funkcję $\varphi: R^2\to R^3$ określoną wzorem $\varphi(x,y)=(x^2y,\ x+y,\ xy^2-1)$. Korzystając z definicji Heinego wykażemy, że funkcja ta ma w punkcie $p_0=(1,-2)$ granicę $g=(-2,-1,3)$.

Niech (p_n) będzie dowolnym ciągiem elemntów przestrzeni R^2 zbieżnym do punktu p_0, o wyrazach różnych od p_0. Ponieważ

$$(p_n=(x_n,y_n)\to(1,-2)) \Leftrightarrow (x_n\to 1 \wedge y_n\to -2),$$

więc

$$\lim_{n\to\infty}\varphi(p_n)=\lim_{n\to\infty}\varphi(x_n,y_n)=\lim_{n\to\infty}(x_n^2y_n,\ x_n+y_n,\ x_ny_n^2-1)=(-2,-1,3).$$

Zgodnie z definicją Heinego oznacza to, że

$$\lim_{(x,y)\to(1,-2)}\varphi(x,y)=(-2,-1,3).$$ ■

PRZYKŁAD 4.2. Niech

$$f(x,y)=\frac{x^2-y^2}{x^2+y^2},\quad (x,y)\in D=R^2-\{(0,0)\}.$$

Wykażemy, że

a) funkcja f ma w punkcie $p_0=(2,-1)$ granicę równą 3/5;

b) funkcja f nie ma granicy w punkcie $p_0=(0,0)$.

a) Niech (p_n) będzie dowolnym ciągiem w przestrzeni R^2 zbieżnym do $p_0=(2,-1)$, którego wyrazy należą do $D-\{p_0\}$. Ponieważ

$$(p_n=(x_n,y_n)\to(2,-1)) \Leftrightarrow (x_n\to 2 \wedge y_n\to -1),$$

więc

$$\lim_{n\to\infty} f(p_n)=\lim_{n\to\infty} f(x_n,y_n)=\lim_{n\to\infty}\frac{x_n^2-y_n^2}{x_n^2+y_n^2}=\frac{4-1}{4+1}=\frac{3}{5}.$$

Z definicji Heinego granicy wynika, że

$$\lim_{(x,y)\to(2,-1)}\frac{x^2-y^2}{x^2+y^2}=\frac{3}{5}.$$

b) Niech

$$p'_n=\left(\frac{1}{n},\frac{1}{n}\right) \quad\text{oraz}\quad p''_n=\left(\frac{1}{n},\frac{3}{n}\right) \quad\text{dla } n\in N.$$

Ciągi (p'_n) i (p''_n) są zbieżne do punktu $p_0=(0,0)$, przy czym $p'_n \neq p_0$ i $p''_n \neq p_0$ dla $n \in N$. Natomiast granice ciągów $(f(p'_n))$ i $(f(p''_n))$ są różne:

$$\lim_{n\to\infty} f(p'_n) = \lim_{n\to\infty} f(\frac{1}{n},\frac{1}{n}) = \lim_{n\to\infty} 0 = 0,$$

$$\lim_{n\to\infty} f(p''_n) = \lim_{n\to\infty} f(\frac{1}{n},\frac{3}{n}) = \lim_{n\to\infty} \frac{-4}{5} = -\frac{4}{5}.$$

Wynika stąd, że funkcja f nie ma granicy w punkcie $p_0=(0,0)$. ■

Podobnie łatwo wykazuje się, że

a) $\lim\limits_{(x,y)\to(1,0)} \dfrac{x-y}{\sqrt{x^2+y^2}} = 1$, b) $\lim\limits_{(x,y)\to(0,0)} \dfrac{2}{x^2+y^2} = +\infty$,

c) $\lim\limits_{(x,y,z)\to(0,1,-1)} \dfrac{xy}{y^2-z} = 0$, d) $\lim\limits_{(x,y)\to(0,0)} \dfrac{x-y}{\sqrt{x^2+y^2}}$ nie istnieje,

e) $\lim\limits_{(x,y)\to(1,-1)} \dfrac{x^2-y^2}{x+y} = 2$, f) $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2}$ nie istnieje.

U w a g a . Zamiast

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y), \qquad \lim_{(x,y,z)\to(x_0,y_0,z_0)} g(x,y,z)$$

piszemy także

$$\lim_{\substack{x\to x_0\\ y\to y_0}} f(x,y), \qquad \lim_{\substack{x\to x_0\\ y\to y_0\\ z\to z_0}} g(x,y,z).$$

GRANICE ITEROWANE. Dla wygody i jasności zapisu ograniczymy się tu do funkcji dwóch zmiennych, chociaż o granicach iterowanych można mówić dla funkcji dowolnej liczby zmiennych.

Granicę funkcji dwóch zmiennych

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)$$

nazywamy g r a n i c ą p o d w ó j n ą (krótko: granicą) w odróżnieniu od g r a n i c i t e r o w a n y c h :

$$\lim_{x\to x_0}(\lim_{y\to y_0} f(x,y)) \quad \text{oraz} \quad \lim_{y\to y_0}(\lim_{x\to x_0} f(x,y)).$$

Granice iterowane funkcji nie muszą być jednakowe, nawet jeżeli obie istnieją (przykład 4.3 a)). Z istnienia granic iterowanych (nawet jednakowych) nie wynika istnienie granicy podwójnej (przykład 4.3 b)), ani też z istnienia granicy podwójnej nie wynika istnienie granic iterowanych.

P R Z Y K Ł A D 4.3.

a) Niech

$$f(x,y) = \frac{x^2-y^2}{x^2+y^2} \quad \text{dla } (x,y) \neq (0,0).$$

Obliczymy granice iterowane tej funkcji w punkcie $p_0=(0,0)$:

$$\lim_{x\to 0}(\lim_{y\to 0} \frac{x^2-y^2}{x^2+y^2}) = \lim_{x\to 0} 1 = 1,$$

$$\lim_{y\to 0}(\lim_{x\to 0} \frac{x^2-y^2}{x^2+y^2}) = \lim_{y\to 0}(-1) = -1.$$

Natomiast

$$\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}$$

nie istnieje, co wykazaliśmy w przykładzie 4.2.

b) Niech

$$f(x,y) = \frac{xy}{x^2+y^2} \quad \text{dla } (x,y) \neq (0,0).$$

Obliczamy granice iterowane tej funkcji w punkcie $p_0=(0,0)$:

$$\lim_{x\to 0}(\lim_{y\to 0} \frac{xy}{x^2+y^2}) = \lim_{x\to 0} 0 = 0,$$

$$\lim_{y\to 0}(\lim_{x\to 0} \frac{xy}{x^2+y^2}) = \lim_{y\to 0} 0 = 0.$$

Granica

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$$

nie istnieje, gdyż dla ciągów (p'_n) i (p''_n) punktów $p'_n=(\frac{1}{n},\frac{1}{n})$ oraz $p''_n=(\frac{1}{n},\frac{2}{n})$, $n \in N$, zbieżnych do $p_0=(0,0)$, ciągi $(f(p'_n))$ i $(f(p''_n))$ mają różne granice:

$$\lim_{n\to\infty} f(p'_n)=\frac{1}{2}, \quad \lim_{n\to\infty} f(p''_n)=\frac{2}{5}. \qquad \blacksquare$$

W dalszym ciągu termin "granica funkcji" będzie zawsze oznaczał granicę w sensie wcześniej przyjętych definicji, czyli granicę podwójną.

CIĄGŁOŚĆ FUNKCJI WIELU ZMIENNYCH. Załóżmy, że dana jest funkcja

$$f: D\to R, \quad \text{przy czym } D\subset R^n.$$

Załóżmy, że p_0 należy do zbioru D i jest punktem skupienia tego zbioru.

Funkcję f nazywamy **ciągłą w punkcie** $p_0\in D$, gdy ma w tym punkcie granicę równą swojej wartości, to znaczy

$$f \text{ jest ciągła w } p_0 \overset{\text{def}}{\Leftrightarrow} \lim_{p\to p_0} f(p)=f(p_0).$$

Jeżeli funkcja nie jest ciągła w punkcie $p_0\in D$, to punkt ten nazywamy **punktem nieciągłości** funkcji.

Mówimy, że funkcja jest **ciągła na zbiorze** $A\subset D$, gdy jest ciągła w każdym punkcie tego zbioru.

Na przykład każda z funkcji

$$f(x,y)=x^2y-y^3, \quad g(x,y)=\sqrt{y-x}, \quad h(x,y)=\frac{xy}{x^2+y^2}$$

jest funkcją ciągłą w dowolnym punkcie swojej dziedziny. Natomiast funkcja określona wzorem

$$F(x,y)=\begin{cases}\sqrt{1-x^2-y^2}, & \text{gdy } x^2+y^2\le 1,\\ 1, & \text{gdy } x^2+y^2>1\end{cases}$$

jest ciągła na całej płaszczyźnie R^2 z wyjątkiem punktów leżących na okręgu $x^2+y^2=1$.

W przypadku gdy funkcja $f(p)=f(x_1,\dots,x_n)$ jest określona na pewnym otoczeniu punktu p_0 i jest ciągła w tym punkcie, to funkcja ta jest ciągła w punkcie p_0 ze względu na każdą ze zmiennych oddzielnie.

Na przykład, jeżeli funkcja $z=f(x,y)$ jest określona na pewnym otoczeniu punktu $p_0=(x_0,y_0)$ i jest w tym punkcie ciągła, to funkcja $\varphi(x)=f(x,y_0)$ jest ciągła w punkcie x_0 oraz funkcja $\psi(y)=f(x_0,y)$ jest ciągła w punkcie y_0.

Jednakże funkcja ciągła ze względu na każdą ze zmiennych oddzielnie nie musi być ciągła jako funkcja wielu zmiennych, o czym świadczy następujący przykład:

PRZYKŁAD 4.4. Funkcja

$$f(x,y)=\begin{cases}\dfrac{xy}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0)\end{cases}$$

jest ciągła w punkcie $p_0=(0,0)$ ze względu na każdą zmienną oddzielnie, gdyż funkcje $\varphi(x)=f(x,0)\equiv 0$ oraz $\psi(y)=f(0,y)\equiv 0$ jako stałe są funkcjami ciągłymi. Jednakże funkcja f nie jest ciągła w punkcie $p_0=(0,0)$ jako funkcja dwóch zmiennych, gdyż (jak to pokazaliśmy w przykładzie 4.3. b)) nie ma granicy w tym punkcie. $\blacksquare$

WŁASNOŚCI FUNKCJI CIĄGŁYCH. W pierwszym tomie tej książki, omawiając ciągłość funkcji jednej zmiennej, podaliśmy trzy ważne własności funkcji ciągłych: własność o lokalnym zachowaniu znaku, własność Darboux i własność Weierstrassa. Obecnie podamy analogiczne twierdzenia dla funkcji ciągłych wielu zmiennych. Załóżmy, że

$$f: D\to R, \quad D\subset R^n.$$

WŁASNOŚĆ I (o lokalnym zachowaniu znaku). Jeżeli funkcja f jest ciągła w punkcie $p_0\in D$ i ma w tym punkcie wartość dodatnią (ujemną), to istnieje takie otoczenie $U(p_0)$ punktu p_0, że funkcja f ma wartości dodatnie (ujemne) w każdym punkcie $p\in U(p_0)\cap D$.

WŁASNOŚĆ II (Darboux - o przyjmowaniu wartości pośrednich). Jeżeli funkcja f jest ciągła na pewnym obszarze, do którego należą punkty p_1 i p_2, to funkcja f przyjmuje na tym obszarze każdą wartość pośrednią między $f(p_1)$ i $f(p_2)$.

W szczególności, gdy $f(p_1)\cdot f(p_2)<0$, to w pewnym punkcie p_0 tego obszaru $f(p_0)=0$.

WŁASNOŚĆ III (Weierstrassa - o osiąganiu najmniejszej i największej wartości). Jeżeli funkcja f jest ciągła na obszarze domkniętym i ograniczonym D, to jest ona ograniczona, przy czym osiąga na tym obszarze swoją wartość najmniejszą i największą, to znaczy

$$\bigvee_{p_1,p_2\in D}\ \bigwedge_{p\in D} f(p_1)\le f(p)\le f(p_2).$$

ZADANIA DO ROZWIĄZANIA.

1. Obliczyć:

a) $\lim\limits_{(x,y)\to(-1,1)}\dfrac{2xy}{x^2+y^2}$, b) $\lim\limits_{(x,y)\to(0,0)}\dfrac{x^2-y^2}{x-y}$,

c) $\lim\limits_{(x,y)\to(-1,2)}\dfrac{x-y}{x+y}$, d) $\lim\limits_{(x,y)\to(0,0)}\dfrac{2-x}{x^2+y^2}$,

e) $\lim\limits_{(x,y)\to(0,1)}\dfrac{x^2}{y^2+1}$, f) $\lim\limits_{(x,y,z)\to(0,0,-2)}\dfrac{z}{x^2+y^2}$.

2. Wykazać, że funkcja $z=f(x,y)$ nie ma granicy w punkcie $p_0=(x_0,y_0)$, gdy:

a) $f(x,y)=\dfrac{x^2}{x^2+y^2}$, $p_0=(0,0)$, b) $f(x,y)=\dfrac{x^2}{y^2-1}$, $p_0=(0,1)$,

c) $f(x,y)=\dfrac{y}{x+1}$, $p_0=(-1,0)$, d) $f(x,y)=\dfrac{xy}{x^2-y^2}$, $p_0=(0,0)$.

3. Obliczyć:

a) $\lim\limits_{x\to0}(\lim\limits_{y\to0}\dfrac{x+y}{x-y})$, $\lim\limits_{y\to0}(\lim\limits_{x\to0}\dfrac{x+y}{x-y})$,

b) $\lim\limits_{x\to0}(\lim\limits_{y\to0}\dfrac{x^2+xy}{x^2+y^2})$, $\lim\limits_{y\to0}(\lim\limits_{x\to0}\dfrac{x^2+xy}{x^2+y^2})$,

c) $\lim\limits_{x\to1}(\lim\limits_{y\to-1}\dfrac{xy+1}{x^2-y^2})$, $\lim\limits_{y\to-1}(\lim\limits_{x\to1}\dfrac{xy+1}{x^2-y^2})$.

4. Zbadać ciągłość funkcji f ze względu na każdą zmienną oddzielnie w punkcie $p_0=(0,0)$, gdy:

a) $f(x,y)=\begin{cases}\dfrac{x^4-y^4}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0),\end{cases}$

b) $f(x,y)=\begin{cases}\dfrac{x^2-y^2}{x^2+y^2}, & (x,y)\neq(0,0),\\ 1, & (x,y)=(0,0),\end{cases}$

c) $f(x,y)=\begin{cases}\dfrac{x^3-xy}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0).\end{cases}$

Wykazać, że funkcja f jako funkcja dwóch zmiennych jest ciągła w punkcie p_0, gdy jest określona wzorem a) oraz nie jest ciągła w p_0, gdy jest określona wzorem b) lub c).

Odpowiedzi.

1. a) -1, b) 0, c) -3, d) $+\infty$, e) 0, f) $-\infty$.
2. Wsk. Obliczyć granicę ciągu $(f(x_n,y_n))$ na przykład dla ciągów:
   a) $(\frac{1}{n},\frac{1}{n})$, $(\frac{1}{n},\frac{2}{n})$, b) $(\frac{1}{n},1+\frac{1}{n})$, $(\frac{1}{n},1+\frac{1}{n^2})$,
   c) $(-1+\frac{1}{n},\frac{1}{n})$, $(-1+\frac{1}{n},\frac{2}{n})$, d) $(\frac{2}{n},\frac{1}{n})$, $(\frac{3}{n},\frac{1}{n})$.
3. a) 1; -1; b) 1; 0; c) $-1/2$; $1/2$.
4. a) ciągła ze względu na każdą ze zmiennych w punkcie (0,0) gdyż $f(x,0)=x^2$, $x\in R$ oraz $f(0,y)=-y^2$, $y\in R$;
   b) ciągła ze względu na x, gdyż $f(x,0)=1$, $x\in R$; nie jest ciągła ze względu na y, gdyż $f(0,y)=-1$ dla $y\neq0$ i $f(0,y)=1$ dla $y=0$;
   c) ciągła ze względu na każdą ze zmiennych, gdyż $f(x,0)=x$, $x\in R$ oraz $f(0,y)=0$, $y\in R$.

5. POCHODNE CZĄSTKOWE. RÓŻNICZKA FUNKCJI.

POCHODNE CZĄSTKOWE. Załóżmy, że funkcja $f: D \to R$, $D \subset R^2$, jest określona na pewnym otoczeniu $U(p_0)$ punktu $p_0 = (x_0, y_0)$. Niech $\varphi_1(x) = f(x, y_0)$ oraz $\varphi_2(y) = f(x_0, y)$.

Jeżeli istnieje pochodna funkcji φ_1 w punkcie x_0, to nazywamy ją **pochodną cząstkową funkcji f względem zmiennej x w punkcie** $p_0 = (x_0, y_0)$ i oznaczamy

$$f'_x(x_0, y_0) \quad \text{lub} \quad \frac{\partial f}{\partial x}(x_0, y_0).$$

Jeżeli istnieje pochodna funkcji φ_2 w punkcie y_0, to nazywamy ją **pochodną cząstkową funkcji f względem zmiennej y w punkcie** $p_0 = (x_0, y_0)$ i oznaczamy

$$f'_y(x_0, y_0) \quad \text{lub} \quad \frac{\partial f}{\partial y}(x_0, y_0).$$

Stąd i z definicji pochodnej funkcji jednej zmiennej wynika, że

$$f'_x(x_0, y_0) = \varphi'_1(x_0) = \lim_{\Delta x \to 0} \frac{\varphi_1(x_0 + \Delta x) - \varphi_1(x_0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}$$

i analogicznie

$$f'_y(x_0, y_0) = \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}.$$

lub (przy innych oznaczeniach):

$$f'_x(x_0, y_0) = \lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0},$$

$$f'_y(x_0, y_0) = \lim_{y \to y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0}.$$

Pochodne cząstkowe funkcji w punkcie są, zgodnie z definicją, pewnymi liczbami. Jeżeli funkcja f ma pochodne cząstkowe w każdym punkcie $p = (x, y) \in A \subset D$, to funkcje

$f'_x: (x, y) \to f'_x(x, y)$, $(x, y) \in A$ oraz $f'_y: (x, y) \to f'_y(x, y)$, $(x, y) \in A$

nazywamy **pochodnymi cząstkowymi funkcji f na zbiorze** A odpowiednio względem zmiennej x lub zmiennej y i oznaczamy odpowiednio

$$f'_x \text{ lub } \frac{\partial f}{\partial x} \quad \text{oraz} \quad f'_y \text{ lub } \frac{\partial f}{\partial y}.$$

Analogicznie definiujemy pochodne cząstkowe funkcji n zmiennych, gdy $n > 2$. Załóżmy zatem, że funkcja $f: D \to R$, $D \subset R^n$, jest określona na pewnym otoczeniu $U(p_0)$ punktu $p_0 = (x_1^0, x_2^0, \ldots, x_n^0)$. Niech φ_k, $k = 1, 2, \ldots, n$, oznacza funkcję jednej zmiennej x_k określoną wzorem

$$\varphi_k(x_k) = f(x_1^0, \ldots, x_{k-1}^0, x_k, x_{k+1}^0, \ldots, x_n^0).$$

Jeżeli istnieje pochodna funkcji φ_k w punkcie x_k^0, to nazywamy ją **pochodną cząstkową funkcji f względem zmiennej** x_k w punkcie $p_0 = (x_1^0, x_2^0, \ldots, x_n^0)$ i oznaczamy

$$f'_{x_k}(x_1^0, \ldots, x_n^0) \quad \text{lub} \quad \frac{\partial f}{\partial x_k}(x_1^0, \ldots, x_n^0),$$

albo też krócej:

$$f'_{x_k}(p_0) \quad \text{lub} \quad \frac{\partial f}{\partial x_k}(p_0).$$

Zatem

$$f'_{x_k}(x_1^0, \ldots, x_n^0) = \lim_{x_k \to x_k^0} \frac{f(x_1^0, \ldots, x_{k-1}^0, x_k, x_{k+1}^0, \ldots, x_n^0) - f(x_1^0, \ldots, x_k^0, \ldots, x_n^0)}{x_k - x_k^0}.$$

Jeżeli funkcja f ma pochodną cząstkową względem zmiennej x_k w każdym punkcie $p = (x_1, \ldots, x_n)$ pewnego zbioru $A \subset D$, to funkcję

$$f'_{x_k}: p \to f'_{x_k}(p), \quad p \in A,$$

nazywamy pochodną cząstkową funkcji f względem zmiennej x_k na zbiorze A.

INTERPRETACJA GEOMETRYCZNA. Niech będzie dana funkcja f dwóch zmiennych x i y określona na pewnym otoczeniu punktu $p_0 = (x_0, y_0)$. Załóżmy, że funkcja f ma pochodne cząstkowe w punkcie p_0. Niech wykresem tej funkcji będzie pewna powierzchnia S. Przetnijmy

tę powierzchnię płaszczyznami π_1: $y = y_0$ i π_2: $x = x_0$, a linie przecięcia tych płaszczyzn z powierzchnią S oznaczmy odpowiednio: K_1 i K_2 (rys. 5.1). Linie K_1 i K_2 są więc określone równaniami

$$K_1: \begin{cases} z = f(x,y), \\ y = y_0, \end{cases} \qquad K_2: \begin{cases} z = f(x,y), \\ x = x_0. \end{cases}$$

W płaszczyźnie π_1 poprowadźmy prostą styczną do krzywej K_1 w punkcie

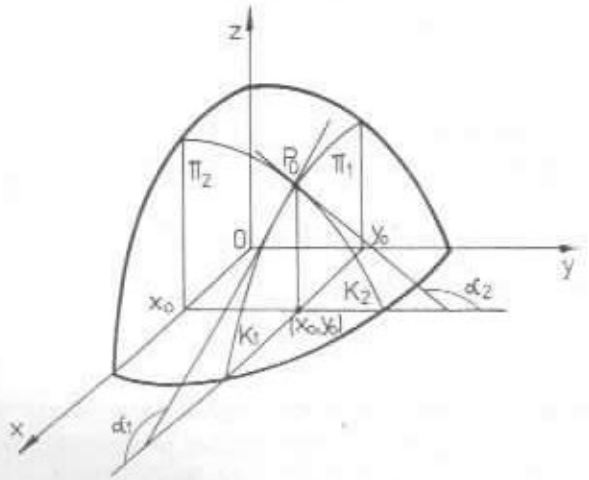
$$P_0 = (x_0, y_0, f(x_0, y_0)).$$

Wartość pochodnej f'_x w punkcie (x_0, y_0) jest równa tangensowi kąta α_1, jaki tworzy styczna do krzywej K_1 w punkcie P_0 z dodatnim kierunkiem osi 0x, czyli

$$f'_x(x_0, y_0) = \operatorname{tg}\alpha_1.$$

Analogicznie, wartość pochodnej f'_y w punkcie (x_0, y_0) jest równa tangensowi kąta α_2, jaki tworzy styczna do krzywej K_2 w punkcie P_0 z dodatnim kierunkiem osi 0y, czyli

$$f'_y(x_0, y_0) = \operatorname{tg}\alpha_2.$$



Rys 5.1.

PRZYKŁADY OBLICZANIA POCHODNYCH CZĄSTKOWYCH. Z przyjętego określenia wynika, że znajdowanie pochodnych cząstkowych funkcji wielu zmiennych sprowadza się do różniczkowania odpowiednich funkcji jednej zmiennej.

P R Z Y K Ł A D 5.1. Niech będzie dana funkcja dwóch zmiennych określona wzorem

(1) $\qquad f(x,y) = x^2 \sin xy, \quad (x,y) \in R^2.$

a) Obliczymy pochodne cząstkowe tej funkcji w punkcie $p_0 = (2,-1)$. Zgodnie z definicją pochodnych cząstkowych

$$f'_x(2,-1) = \varphi'_1(2), \qquad f'_y(2,-1) = \varphi'_2(-1),$$

gdzie $\varphi_1(x) = f(x,-1) = -x^2 \sin x$, $\varphi_2(y) = f(2,y) = 4\sin 2y$. Stąd

$$\varphi'_1(x) = -2x\sin x - x^2\cos x, \qquad \varphi'_2(y) = 8\cos 2y,$$
$$\varphi'_1(2) = -4(\sin 2 + \cos 2), \qquad \varphi'_2(-1) = 8\cos 2,$$
$$f'_x(2,-1) = -4(\sin 2 + \cos 2), \qquad f'_y(2,-1) = 8\cos 2.$$

b) Znajdziemy teraz pochodne cząstkowe funkcji f w dowolnym punkcie $(x,y) \in R^2$.

Aby znaleźć pochodną f'_x, zmienną y we wzorze (1) traktujemy jako stałą. Otrzymujemy

(2) $\qquad f'_x = 2x\sin xy + x^2 y\cos xy.$

Analogicznie, przyjmując w (1), że x jest stałą, otrzymujemy

(3) $\qquad f'_y = x^3\cos xy.$

Wartości pochodnych cząstkowych funkcji f w punkcie $(2,-1)$ obliczone w a) można oczywiście otrzymać z (2) i (3) przyjmując $x = 2$ i $y = -1$. ■

P R Z Y K Ł A D 5.2. Zbadamy istnienie pochodnych cząstkowych funkcji f w punkcie $p_0 = (0,0)$, gdy:

$$\text{a) } f(x,y) = \begin{cases} \dfrac{x^3 - y^2}{x^2 + y^2} & \text{dla } (x,y) \neq (0,0), \\ 0 & \text{dla } (x,y) = (0,0), \end{cases}$$

b) $f(x,y)=\begin{cases}\dfrac{xy}{x^2+y^2} & \text{dla } (x,y)\neq(0,0),\\ 0 & \text{dla } (x,y)=(0,0).\end{cases}$

a) Ponieważ $\varphi_1(x)=f(x,0)=x$ dla każdego $x\in R$, więc $\varphi_1'(x)=1$, $\varphi_1'(0)=1$, a w konsekwencji

$$f_x'(0,0)=1.$$

Natomiast funkcja

$$\varphi_2(y)=f(0,y)=\begin{cases}-1 & \text{dla } y\neq 0,\\ 0 & \text{dla } y=0\end{cases}$$

nie jest ciągła w punkcie $y_0=0$, a więc nie ma pochodnej w tym punkcie, co oznacza, że w punkcie $p_0=(0,0)$ nie istnieje pochodna cząstkowa f_y'.

b) Ponieważ

$$\varphi_1(x)=f(x,0)=0,\ x\in R \quad \text{oraz} \quad \varphi_2(y)=f(0,y)=0,\ y\in R,$$

więc

$$\varphi_1'(x)=0,\quad \varphi_1'(0)=0,\quad f_x'(0,0)=0$$

oraz

$$\varphi_2'(y)=0,\quad \varphi_2'(0)=0,\quad f_y'(0,0)=0.$$

Warto jeszcze zauważyć, że funkcja ta, choć ma obie pochodne cząstkowe w punkcie $p_0=(0,0)$, nie jest ciągła w tym punkcie (gdyż nie ma granicy, co pokazaliśmy w przykładzie 4.3.). ■

Uwaga. Z powyższego przykładu wynika, że dla funkcji wielu zmiennych istnienie pochodnych cząstkowych nie gwarantuje ciągłości funkcji. Przypomnijmy, że dla funkcji jednej zmiennej z istnienia pochodnej w pewnym punkcie wynika ciągłość funkcji w tym punkcie. Z istnienia pochodnych cząstkowych wynika jedynie ciągłość funkcji ze względu na każdą zmienną oddzielnie. Dowodzi się jednak, że funkcja, która ma ciągłe pochodne cząstkowe na pewnym obszarze, jest na tym obszarze funkcją ciągłą.

PRZYKŁAD 5.3. Obliczymy pochodne cząstkowe funkcji:

a) $z(x,y)=x^2y-y^5+2x$,

b) $w(x,y)=(x-2y)^3+x^3$,

c) $u(x,z)=\ln(x^2+z^4+1)$,

d) $t(y,z)=y\sqrt{2+z^2}-z$,

e) $p(x,y,z)=xe^{-2y}+z^2e^{-2x}$,

f) $r(x,y,z)=\operatorname{arctg}\dfrac{y}{x}-\operatorname{arctg}\dfrac{1}{z}$.

Obliczamy:

a) $z_x'=2xy+2$, $z_y'=x^2-5y^4$;

b) $w_x'=3(x-2y)^2+3x^2$, $w_y'=-6(x-2y)^2$;

c) $u_x'=\dfrac{2x}{x^2+z^4+1}$, $u_z'=\dfrac{4z^3}{x^2+z^4+1}$;

d) $t_y'=\sqrt{2+z^2}$, $t_z'=\dfrac{zy}{\sqrt{2+z^2}}-1$;

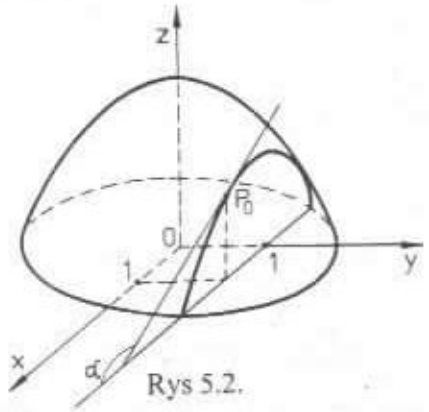
e) $p_x'=e^{-2y}-2z^2e^{-2x}$, $p_y'=-2xe^{-2y}$, $p_z'=2ze^{-2x}$;

f) $r_x'=\dfrac{-y}{x^2+y^2}$, $r_y'=\dfrac{x}{x^2+y^2}$, $r_z'=\dfrac{1}{1+z^2}$. ■

PRZYKŁAD 5.4. Obliczymy kąt, jaki tworzy z dodatnim kierunkiem osi 0x styczna do linii

$$2z=4-x^2-y^2,\quad y=1$$

w punkcie $P_0=(1,1,1)$.



Rys 5.2.

Powierzchnia o równaniu $2z=4-x^2-y^2$ jest paraboloidą obrotową (rys 5.2.), a opisaną linię otrzymujemy przecinając tę powierzchnię płaszczyzną $y=1$. Tangens szukanego kąta α jest równy wartości pochodnej z_x' w punkcie $(1,1)$, $\operatorname{tg}\alpha=z_x'(1,1)$. Obliczamy: $z_x'(x,y)=-x$, $z_x'(1,1)=-1$. Tak więc

$$\operatorname{tg}\alpha=-1,\quad \alpha=3\pi/4.$$ ■

POCHODNE CZĄSTKOWE WYŻSZYCH RZĘDÓW. Pochodne cząstkowe funkcji wielu zmiennych nazywane są też pochodnymi cząstkowymi pierwszego rzędu w odróżnieniu od pochodnych cząstkowych wyższych rzędów, o których mówimy niżej.

Pochodne cząstkowe funkcji n zmiennych są również funkcjami tych zmiennych i można mówić o ich pochodnych cząstkowych. Dla przykładu: niech $z = x^2y - 3y^2$, $(x,y) \in R^2$, wówczas $z'_x = 2xy$, $z'_y = x^2 - 6y$ oraz $(z'_x)'_x = 2y$, $(z'_x)'_y = 2x$, $(z'_y)'_x = 2x$, $(z'_y)'_y = -6$ dla $(x,y) \in R^2$.

Pochodne cząstkowe pochodnych cząstkowych funkcji f nazywamy **pochodnymi cząstkowymi drugiego rzędu** funkcji f, przy czym przyjmujemy oznaczenia

$$\frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_i}\right) \equiv \frac{\partial^2 f}{\partial x_k \partial x_i} \quad \text{lub} \quad (f'_{x_i})'_{x_k} \equiv f''_{x_i x_k}.$$

W szczególności dla funkcji dwóch zmiennych $z = f(x,y)$ mamy

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \equiv \frac{\partial^2 f}{\partial x^2}, \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \equiv \frac{\partial^2 f}{\partial y \partial x},$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \equiv \frac{\partial^2 f}{\partial x \partial y}, \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \equiv \frac{\partial^2 f}{\partial y^2}$$

lub

$$(f'_x)'_x \equiv f''_{xx}, \quad (f'_x)'_y \equiv f''_{xy}, \quad (f'_y)'_x \equiv f''_{yx}, \quad (f'_y)'_y \equiv f''_{yy}.$$

Funkcja n zmiennych może mieć n^2 pochodnych cząstkowych rzędu drugiego, przy czym pochodne

$$\frac{\partial^2 f}{\partial x_k \partial x_i}, \quad \text{gdy } i \neq k$$

nazywamy pochodnymi cząstkowymi mieszanymi rzędu drugiego.

Analogicznie definiujemy pochodne cząstkowe wyższych rzędów:

Pochodne cząstkowe rzędu m funkcji f są to pochodne cząstkowe pochodnych cząstkowych rzędu $m-1$ funkcji f.

Oznaczenia pochodnych cząstkowych rzędu m są analogiczne do stosowanych wyżej w przypadku $m = 2$. Dla przykładu

$$\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right) \equiv \frac{\partial^3 f}{\partial x^2 \partial y}, \qquad \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x^2}\right) \equiv \frac{\partial^3 f}{\partial y \partial x^2},$$

$$\frac{\partial}{\partial x}\left(\frac{\partial^3 f}{\partial x \partial y^2}\right) \equiv \frac{\partial^4 f}{\partial x^2 \partial y^2}, \qquad \frac{\partial}{\partial x}\left(\frac{\partial^3 f}{\partial x^3}\right) \equiv \frac{\partial^4 f}{\partial x^4}$$

lub odpowiednio:

$$(f''_{yx})'_x \equiv f'''_{yxx} \equiv f'''_{yx^2}, \qquad (f''_{xx})'_y \equiv f'''_{xxy} \equiv f'''_{x^2y},$$

$$(f'''_{y^2x})'_x \equiv f^{(4)}_{y^2x^2}, \qquad (f'''_{x^3})'_x \equiv f^{(4)}_{x^4}.$$

PRZYKŁAD 5.5. Znajdziemy pochodne cząstkowe drugiego rzędu funkcji określonych wzorami:

a) $f(x,y) = x^2y - e^{-y}$, b) $g(x,y,z) = x^2y^2 - y^2z + 3z^2$.

Obliczamy

a) $f'_x = 2xy$, $f'_y = x^2 + e^{-y}$,

$f''_{xx} = 2y$, $f''_{xy} = 2x$, $f''_{yx} = 2x$, $f''_{yy} = -e^{-y}$.

b) $g'_x = 2xy^2$, $g'_y = 2x^2y - 2yz$, $g'_z = -y^2 + 6z$,

$g''_{xx} = 2y^2$, $g''_{xy} = 4xy$, $g''_{xz} = 0$,

$g''_{yx} = 4xy$, $g''_{yy} = 2x^2 - 2z$, $g''_{yz} = -2y$,

$g''_{zx} = 0$, $g''_{zy} = -2y$, $g''_{zz} = 6$. ■

Zauważmy, że w powyższym przykładzie otrzymaliśmy

$$f''_{xy} = f''_{yx} \quad \text{oraz} \quad g''_{xy} = g''_{yx}, \quad g''_{xz} = g''_{zx}, \quad g''_{yz} = g''_{zy}.$$

Zwróćmy uwagę na to, że pochodne cząstkowe, które okazały się równe, są pochodnymi cząstkowymi mieszanymi rzędu drugiego, różniącymi się tylko kolejnością różniczkowania względem zmiennych. Przykład ten jest ilustracją twierdzenia, które podajemy niżej:

TWIERDZENIE 5.1 (Schwarza). Jeżeli funkcja f zmiennych $x_1, \ldots, x_n$ ma na pewnym otoczeniu punktu $p_0 = (x_1^0, \ldots, x_n^0)$ pochodne cząstkowe mieszane

$$\frac{\partial^2 f}{\partial x_k \partial x_i}, \qquad \frac{\partial^2 f}{\partial x_i \partial x_k}$$

i są one ciągłe w punkcie p_0, to pochodne te są równe w punkcie p_0.

FUNKCJE KLASY C^m. Podobnie jak dla funkcji jednej zmiennej, dla funkcji wielu zmiennych przyjmujemy następujące określenia:

Mówimy, że **funkcja f jest klasy C^0** na pewnym obszarze, gdy jest funkcją ciągłą na tym obszarze.

Mówimy, że **funkcja f jest klasy C^m** na pewnym obszarze, gdy funkcja ta ma ciągłe pochodne cząstkowe rzędu m na tym obszarze.

Mówimy, że **funkcja jest klasy C^∞** na pewnym obszarze, gdy funkcja ta ma ciągłe pochodne cząstkowe dowolnego rzędu na tym obszarze.

Wykazuje się, że $C^0 \supset C^1 \supset C^2 \supset \dots$. Stąd wynika, że funkcja klasy C^m na pewnym obszarze jest na tym obszarze ciągła i ma ciągłe pochodne cząstkowe do m-tego rzędu włącznie.

Łatwo widać, że funkcja $f(x,y)=x^2y^3-2y^4+2$ jest klasy C^∞ na płaszczyźnie R^2, natomiast funkcja $g(x,y)=y^2\sqrt[3]{y}-x^3$ jest klasy C^2 na płaszczyźnie R^2, ale nie jest funkcją klasy C^3 na R^2, gdyż pochodna cząstkowa trzeciego rzędu $g'''_{yyy}(0,0)$ nie istnieje.

POCHODNE CZĄSTKOWE FUNKCJI ZŁOŻONEJ. Weźmy pod uwagę funkcję złożoną $F(p)=f(\varphi(p))$, gdzie

$$\varphi: D\to\Delta,\ D\subset R^m,\ \Delta\subset R^n \qquad \text{oraz} \qquad f:\Delta\to R.$$

Funkcja F jest funkcją m zmiennych rzeczywistych i przyjmuje wartości rzeczywiste.

Jeżeli

$$f(u)=f(u_1,\dots,u_n), \qquad u=(u_1,\dots,u_n)\in\Delta, \tag{5.1}$$

oraz

$$u=\varphi(p)=(u_1(p),\dots,u_n(p)), \qquad p=(x_1,\dots,x_m)\in D, \tag{5.2}$$

to funkcję złożoną $F(p)=f(\varphi(p))$ można zapisać w postaci

$$F(p)=f(u_1(p),\dots,u_n(p)), \qquad p\in D,$$

lub

$$F(x_1,\dots,x_m)=f(u_1(x_1,\dots,x_m),\dots,u_n(x_1,\dots,x_m)),\ (x_1,\dots,x_m)\in D. \tag{5.3}$$

Niżej podajemy twierdzenie o pochodnych cząstkowych funkcji złożonej (5.3).

TWIERDZENIE 5.2. Jeżeli funkcja f dana wzorem (5.1) jest klasy C^1 na obszarze Δ oraz funkcje $u_k(x_1,\dots,x_m)$, $k=1,\dots,n$, występujące w (5.2) mają pochodne cząstkowe na obszarze D, to funkcja złożona F dana wzorem (5.3) ma pochodne cząstkowe na obszarze D, przy czym

$$\frac{\partial F}{\partial x_i}=\sum_{k=1}^{n}\frac{\partial f}{\partial u_k}\frac{\partial u_k}{\partial x_i}, \qquad i=1,\dots,m. \tag{5.4}$$

W szczególności, gdy

$$z=f(u,v), \quad u=u(x,y), \quad v=v(x,y),$$

funkcja złożona $F(x,y)=f(u(x,y),v(x,y))$ jest funkcją dwóch zmiennych x i y. Wzory (5.4) przyjmują wtedy postać

$$\frac{\partial F}{\partial x}=\frac{\partial f}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial f}{\partial v}\frac{\partial v}{\partial x}, \quad \frac{\partial F}{\partial y}=\frac{\partial f}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial f}{\partial v}\frac{\partial v}{\partial y}. \tag{5.5}$$

Rozważmy jeszcze dwa szczególne przypadki:

1) Jeżeli

$$z=f(u,v), \quad u=u(x), \quad v=v(x),$$

to funkcja złożona $F(x)=f(u(x),v(x))$ jest funkcją zmiennej x. Wówczas z (5.4) otrzymujemy

$$\frac{dF}{dx}=\frac{\partial f}{\partial u}\frac{du}{dx}+\frac{\partial f}{\partial v}\frac{dv}{dx} \tag{5.6}$$

(symbole pochodnych cząstkowych w przypadku jednej zmiennej zostały zastąpione symbolami pochodnych zwyczajnych).

2) Jeżeli

$$z=f(x,y), \quad y=y(x),$$

to funkcja złożona $F(x)=f(x,y(x))$ jest funkcją tylko zmiennej x. Uwzględniając stosowane w tym przypadku oznaczenia, ze wzoru (5.6) otrzymujemy

$$\frac{dF}{dx}=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\frac{dy}{dx}.$$

RÓŻNICZKA FUNKCJI DWÓCH ZMIENNYCH. Załóżmy, że funkcja f dwóch zmiennych x i y jest określona na pewnym otoczeniu $U(p_0)$ punktu $p_0=(x_0,y_0)$. Załóżmy, że Δx i Δy są takimi przyrostami zmiennych x i y, że $p=(x_0+\Delta x, y_0+\Delta y)\in U(p_0)$, przy czym $(\Delta x)^2+(\Delta y)^2>0$. Różnicę

$$\Delta f = f(x_0+\Delta x, y_0+\Delta y)-f(x_0,y_0)\equiv f(p)-f(p_0)$$

nazywamy **przyrostem funkcji** f między punktami p i p_0.

Załóżmy, że funkcja f ma ciągłe pochodne cząstkowe w dowolnym punkcie $p\in U(p_0)$.

Różniczką zupełną (krócej: różniczką) funkcji f w punkcie $p_0=(x_0,y_0)$ dla przyrostu $\Delta p=(\Delta x,\Delta y)$ nazywamy wyrażenie

$$df(p_0)\overset{def}{=} f'_x(p_0)\Delta x+f'_y(p_0)\Delta y.$$

Przyrosty Δx i Δy są często oznaczane symbolami dx i dy, a różniczka - skróconym symbolem df:

$$df = f'_x\,dx+f'_y\,dy.$$

Na przykład dla funkcji $z(x,y)=x^2e^{-xy}$ w dowolnym punkcie $(x,y)\in R^2$ mamy:

$$dz=(2x-x^2y)e^{-xy}dx-x^3e^{-xy}dy.$$

TWIERDZENIE 5.3. Jeżeli funkcja f jest klasy C^1 na otoczeniu $U(p_0)$ punktu $p_0=(x_0,y_0)$ oraz przyrosty dx i dy są takie, że $(x_0+dx,y_0+dy)\in U(p_0)$, to

$$\lim_{\substack{dx\to 0\\ dy\to 0}}\frac{\Delta f-df}{\sqrt{(dx)^2+(dy)^2}}=0.$$

Twierdzenie to pozwala dla małych przyrostów dx i dy przyrost Δf funkcji z dużym przybliżeniem zastąpić różniczką df. Stąd otrzymujemy wzór przybliżony

$$\Delta f\approx df$$

i w konsekwencji

(5.7) $$f(x_0+dx,y_0+dy)\approx f(x_0,y_0)+f'_x(x_0,y_0)dx+f'_y(x_0,y_0)dy.$$

Oczywiście przybliżenie będzie tym lepsze, im mniejsze będą przyrosty dx i dy. Co więcej, z twierdzenia 5.3 wynika, że błąd jaki popełniamy stosując wzór (5.7) dąży do zera szybciej niż odległość $\rho=\sqrt{(dx)^2+(dy)^2}$ punktów (x_0,y_0) i (x_0+dx,y_0+dy).

PRZYKŁAD 5.6. Stosując wzór (5.7) obliczymy przybliżoną wartość $\sqrt{(1,12)^2+3,2}$.

Niech $f(x,y)=\sqrt{x^2+y}$, $x_0=1$, $y_0=3$, $dx=0,12$, $dy=0,20$.

Wówczas $$df=\frac{x}{\sqrt{x^2+y}}dx+\frac{1}{2\sqrt{x^2+y}}dy,$$

$$f(1,3)=2,\quad df=0,5\cdot 0,12+0,25\cdot 0,2=0,11.$$

Zatem $\sqrt{(1,12)^2+3,2}\approx 2+0,11=2,11$. ■

Załóżmy, że funkcja f zmiennych x i y jest klasy C^2 na pewnym obszarze $D\subset R^2$. Wówczas, przy ustalonych przyrostach dx i dy, jej różniczka

$$df=f'_x(x,y)dx+f'_y(x,y)dy,\quad p=(x,y)\in D$$

jest również funkcją zmiennych x i y na obszarze D.

Różniczkę różniczki df, przy ustalonych przyrostach dx i dy nazywamy **różniczką drugiego rzędu** funkcji f i oznaczamy $d^2f(p)$ lub krócej d^2f. Zatem

$$\begin{aligned}d^2f=d(df)=d(f'_x dx+f'_y dy)=\\ =(f''_{xx}dx+f''_{yx}dy)dx+(f''_{xy}dx+f''_{yy}dy)dy=\\ =f''_{xx}(dx)^2+2f''_{xy}dxdy+f''_{yy}(dy)^2.\end{aligned}$$

Ostatni wzór przypomina niewątpliwie znany z algebry wzór $(a+b)^2=a^2+2ab+b^2$ i zapisujemy go w sposób umowny w postaci

$$d^2f=(f'_x dx+f'_y dy)^{(2)}$$

lub

$$d^2f=\left(\frac{\partial}{\partial x}dx+\frac{\partial}{\partial y}dy\right)^{(2)}f.$$

Przyjmując oznaczenia: $(dx)^2=dx^2$, $(dy)^2=dy^2$ otrzymujemy

$$d^2f = f''_{xx}dx^2 + 2f''_{xy}dxdy + f''_{yy}dy^2.$$

Analogicznie określamy różniczki wyższych rzędów:

$$d^nf = d(d^{n-1}f), \quad n = 2,3,\dots .$$

Przy założeniu, że funkcja f jest klasy C^n na obszarze D, a przyrosty dx i dy są ustalone, można udowodnić indukcyjnie, że

$$d^nf = \frac{\partial^n f}{\partial x^n}dx^n + \binom{n}{1}\frac{\partial^n f}{\partial x^{n-1}\partial y}dx^{n-1}dy + \dots + \binom{n}{n}\frac{\partial^n f}{\partial y^n}dy^n.$$

Wzór ten łatwo zapamiętać, gdy zapiszemy go w sposób umowny w postaci:

$$d^nf = (\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy)^{(n)}f.$$

Wprowadzone pojęcie różniczki funkcji dwóch zmiennych przenosi się łatwo na przypadek funkcji większej liczby zmiennych. Czytelników zainteresowanych tym tematem odsyłamy do obszerniejszych podręczników.

WZÓR TAYLORA. Wzór Taylora omawialiśmy już w pierwszym tomie (rozdz. III. 4) dla funkcji jednej zmiennej. Obecnie podamy uogólnienie tego wzoru na przypadek funkcji dwóch zmiennych (przy odpowiednio zmienionych założeniach wzór ten może być zapisany w tej postaci dla funkcji większej liczby zmiennych).

TWIERDZENIE 5.4. (Taylora). Jeżeli funkcja f zmiennych x i y jest klasy C^n na pewnym otoczeniu $U(p_0)$ punktu $p_0 = (x_0, y_0)$ i punkt $p = (x_0 + dx, y_0 + dy)$ jest dowolnym punktem tego otoczenia, to istnieje liczba $\theta \in (0,1)$ taka, że

$$(5.8) \qquad f(p) - f(p_0) = \frac{df(p_0)}{1!} + \frac{d^2f(p_0)}{2!} + \dots + \frac{d^{n-1}f(p_0)}{(n-1)!} + R_n,$$

gdzie $R_n = \frac{1}{n!}d^nf(p_0 + \theta dp)$ oraz $p_0 + \theta dp = (x_0 + \theta dx, y_0 + \theta dy)$.

Na przykład dla $n = 2$ i przy oznaczeniach $dx = h$, $dy = k$ wzór (5.8) przyjmuje postać

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = f'_x(x_0, y_0)h + f'_y(x_0, y_0)k + R_2,$$

gdzie

$$R_2 = \frac{1}{2}d^2f(x_0 + \theta h, y_0 + \theta k) = \frac{1}{2}(f''_{xx}h^2 + 2f''_{xy}hk + f''_{yy}k^2),$$

przy czym pochodne f''_{xx}, f''_{xy}, f''_{yy} obliczone są w punkcie $(x_0 + \theta h, y_0 + \theta k)$, a θ oznacza pewną liczbę z przedziału $(0,1)$.

ZADANIA DO ROZWIĄZANIA.

1. Znaleźć pochodne cząstkowe pierwszego rzędu następujących funkcji:
   a) $f(x,y) = x^2\cos y + 3y^2\sin x$, b) $f(x,y) = xe^{-2y} + y^2$,
   c) $f(x,y) = \arcsin\frac{y}{x}$, d) $f(x,y) = \operatorname{arctg}\frac{x+y}{x-y}$,
   e) $f(x,y) = \frac{x^2}{x^2 - y^2}$, f) $f(x,y) = \frac{\ln(1-2y)}{xy}$,
   g) $f(x,y) = y\sqrt{y^2 - x^2}$, h) $f(x,y) = x^y - \ln x$,
   i) $f(x,y,z) = xe^{-z} + y^3z$, j) $f(x,y,z) = z\operatorname{arctg}\frac{y}{x}$.
2. Obliczyć pochodne cząstkowe pierwszego rzędu funkcji f w punkcie p_0, gdy:
   a) $f(x,y) = x\sqrt{y^2 - 3x}$, $p_0 = (3,5)$,
   b) $f(x,y) = x\ln(2y-1) - y^2$, $p_0 = (-1,1)$,
   c) $f(x,y) = x^y - x\ln y$, $p_0 = (1,1)$,
   d) $f(x,y,z) = xe^{-2y} + ye^{-2z} + e^x$, $p_0 = (0,0,0)$.
3. Znaleźć pochodne cząstkowe drugiego rzędu następujących funkcji:
   a) $f(x,y) = \frac{1}{x}\ln(1-y^2)$, b) $f(x,y) = x^2 - \operatorname{arctg}\frac{1}{y}$,
   c) $f(x,y) = (y - x^2)e^{-y}$, d) $f(x,y) = xe^{-xy}$.
4. Obliczyć, o ile istnieją, pochodne $f'_x(0,0)$ i $f'_y(0,0)$, gdy
   a) $f(x,y) = \sqrt{x^2 + y^4}$, b) $f(x,y) = x\sqrt{x^2 + y^2}$,

c) $f(x,y)=\begin{cases}\dfrac{x^2-y^3}{x^2+y^2}, & (x,y)\neq(0,0),\\ 1, & (x,y)=(0,0),\end{cases}$

d) $f(x,y)=\begin{cases}\dfrac{x^3-y^3}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0),\end{cases}$

e) $f(x,y)=\begin{cases}\dfrac{x^2-xy^3+2y^3+y^2}{x^2+y^2}, & (x,y)\neq(0,0),\\ 1, & (x,y)=(0,0).\end{cases}$

5. Wyznaczyć f'_x i f'_y dla dowolnych $(x,y)\in R^2$, gdy

$$f(x,y)=\begin{cases}\dfrac{x^3-y^3}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0).\end{cases}$$

6. Wykazać, że funkcja

$$f(x,y)=\frac{1}{y^2-x^2} \quad \text{dla} \quad y^2-x^2\neq 0$$

spełnia równanie różniczkowe: $y\dfrac{\partial f}{\partial x}+x\dfrac{\partial f}{\partial y}=0$.

7. Sprawdzić, że funkcja

$$f(x,y)=\begin{cases}\dfrac{xy}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0)\end{cases}$$

spełnia równanie $x\dfrac{\partial f}{\partial x}+y\dfrac{\partial f}{\partial y}=0$ dla każdego $(x,y)\in R^2$.

8. Sprawdzić, że funkcja

$$f(x,y)=\frac{x}{x^2+y^2}$$

spełnia tzw. równanie Laplace'a: $\dfrac{\partial^2 f}{\partial x^2}+\dfrac{\partial^2 f}{\partial y^2}=0$ dla każdego $(x,y)\in R^2-\{(0,0)\}$.

9. Korzystając z wyników zadania 5 obliczyć, o ile istnieją, pochodne f''_{xx}, f''_{xy}, f''_{yy} w punkcie $p_0=(0,0)$ funkcji:

$$f(x,y)=\begin{cases}\dfrac{x^3-y^3}{x^2+y^2}, & (x,y)\neq(0,0),\\ 0, & (x,y)=(0,0).\end{cases}$$

Odpowiedzi.

1. a) $f'_x=2x\cos y+3y^2\cos x$, $f'_y=-x^2\sin y+6y\sin x$, b) $f'_x=e^{-2y}$, $f'_y=-2xe^{-2y}+2y$, c) $f'_x=\dfrac{-y}{|x|\sqrt{x^2-y^2}}$, $f'_y=\dfrac{|x|}{x\sqrt{x^2-y^2}}$, d) $f'_x=\dfrac{-y}{x^2+y^2}$, $f'_y=\dfrac{x}{x^2+y^2}$, e) $f'_x=-\dfrac{2xy^2}{(x^2-y^2)^2}$, $f'_y=\dfrac{2x^2y}{(x^2-y^2)^2}$, f) $f'_x=-\dfrac{\ln(1-2y)}{x^2y}$, $f'_y=\dfrac{-2}{(1-2y)xy}-\dfrac{\ln(1-2y)}{xy^2}$, g) $f'_x=\dfrac{-xy}{\sqrt{y^2-x^2}}$, $f'_y=\dfrac{2y^2-x^2}{\sqrt{y^2-x^2}}$, h) $f'_x=yx^{y-1}-x^{-1}$, $f'_y=x^y\ln x$, i) $f'_x=e^{-x}$, $f'_y=3y^2z$, $f'_z=y^3-xe^{-x}$, j) $f'_x=\dfrac{-zy}{x^2+y^2}$, $f'_y=\dfrac{zx}{x^2+y^2}$, $f'_z=\operatorname{arctg}\dfrac{y}{x}$.
2. a) $f'_x(3,5)=23/8$, $f'_y(3,5)=15/4$, b) $f'_x(-1,1)=0$, $f'_y(-1,1)=-4$, c) $f'_x(1,1)=1$, $f'_y(1,1)=-1$, d) $f'_x(0,0,0)=2$, $f'_y(0,0,0)=1$, $f'_z(0,0,0)=0$.
3. a) $f''_{xx}=\dfrac{2}{x^3}\ln(1-y^2)$, $f''_{xy}=\dfrac{2y}{x^2(1-y^2)}$, $f''_{yy}=\dfrac{-2(1+y^2)}{x(1-y^2)^2}$, b) $f''_{xx}=2$, $f''_{xy}=0$, $f''_{yy}=\dfrac{-2y}{(y^2+1)^2}$, c) $f''_{xx}=-2e^{-y}$, $f''_{xy}=2xe^{-y}$, $f''_{yy}=(-2+y-x^2)e^{-y}$, d) $f''_{xx}=y(xy-2)e^{-xy}$, $f''_{xy}=x(xy-2)e^{-xy}$, $f''_{yy}=x^3e^{-xy}$.
4. a) $f'_x(0,0)$ nie istnieje, $f'_y(0,0)=0$; b) $f'_x(0,0)=0$, $f'_y(0,0)=0$; c) $f'_x(0,0)=0$, $f'_y(0,0)$ nie istnieje; d) $f'_x(0,0)=1$, $f'_y(0,0)=-1$; e) $f'_x(0,0)=0$, $f'_y(0,0)=2$.
5. $f'_x(x,y)=\dfrac{x(x^3+3xy^2+2y^3)}{(x^2+y^2)^2}$ dla $(x,y)\neq(0,0)$, $f'_x(0,0)=1$, $f'_y(x,y)=\dfrac{y(-3yx^2-y^3-2x^3)}{(x^2+y^2)^2}$ dla $(x,y)\neq(0,0)$, $f'_y(0,0)=-1$.
9. $f''_{xx}(0,0)=0$, $f''_{yy}(0,0)=0$, $f''_{xy}(0,0)$ nie istnieje, $f''_{yx}(0,0)$ nie istnieje.

6. EKSTREMA FUNKCJI DWÓCH ZMIENNYCH.

EKSTREMA FUNKCJI. Załóżmy, że funkcja f zmiennych x i y jest określona na pewnym otoczeniu punktu $p_0 = (x_0, y_0)$. Przyjmijmy następujące określenia:

Mówimy, że funkcja f ma w punkcie p_0 **maksimum lokalne**, gdy dla punktów p z pewnego otoczenia punktu p_0 zachodzi nierówność $f(p) \le f(p_0)$, czyli

$$(\text{f ma maksimum lokalne w } p_0) \overset{\text{def}}{\Leftrightarrow} \bigvee_{U(p_0)} \bigwedge_{p \in U(p_0)} f(p) \le f(p_0).$$

Analogicznie definiujemy **minimum lokalne**:

$$(\text{f ma minimum lokalne w } p_0) \overset{\text{def}}{\Leftrightarrow} \bigvee_{U(p_0)} \bigwedge_{p \in U(p_0)} f(p) \ge f(p_0).$$

W szczególności, jeżeli dla każdego $p \in U(p_0) - \{p_0\}$ zachodzi nierówność ostra: $f(p) < f(p_0)$ lub $f(p) > f(p_0)$, to mówimy o **maksimum** lub **minimum lokalnym właściwym** funkcji f w punkcie p_0.

Jeżeli funkcja f jest określona na pewnym zbiorze $D \subset R^2$, przy czym $p_0 \in D$ oraz

$$\bigwedge_{p \in D} f(p) \le f(p_0) \quad \text{lub} \quad \bigwedge_{p \in D} f(p) \ge f(p_0),$$

to mówimy, że funkcja f ma w punkcie p_0 odpowiednio **maksimum absolutne** lub **minimum absolutne** na zbiorze D.

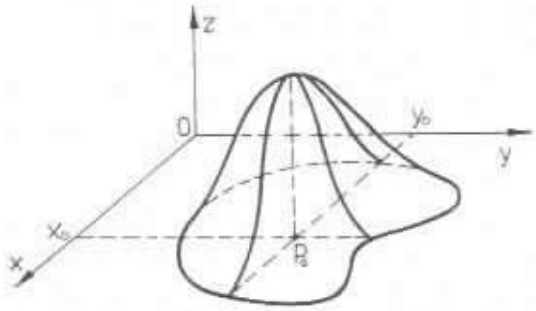
Tak więc maksimum (minimum) absolutne oznacza po prostu największą (najmniejszą) wartość funkcji na rozpatrywanym zbiorze.

Wykazuje się, że funkcja ciągła na obszarze d o m k n i ę t y m i o g r a n i c z o n y m osiąga na tym obszarze minimum i maksimum absolutne (własność Weierstrassa funkcji ciągłych).

Na rysunku 6.1. przedstawiony jest wykres funkcji mającej w punkcie p_0 maksimum lokalne, które jest jednocześnie maksimum absolutnym tej funkcji.

Oczywiście nie zawsze ekstremum lokalne jest ekstremum absolutnym, ani też odwrotnie - ekstremum absolutne nie musi być ekstremum lokalnym. Jeżeli jednak ekstremum absolutne jest osiągane w

punkcie wewnętrznym dziedziny, to jest ono jednocześnie ekstremum lokalnym.



Rys 6.1.

WARUNEK KONIECZNY ISTNIENIA EKSTREMUM. Załóżmy, że funkcja f zmiennych x i y jest określona na pewnym otoczeniu punktu $p_0 = (x_0, y_0)$.

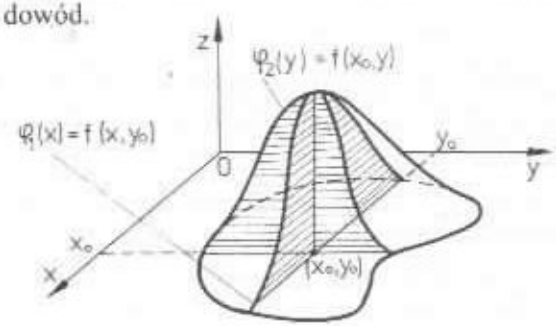
TWIERDZENIE 6.1. Jeżeli funkcja f ma w punkcie (x_0, y_0) pochodne cząstkowe i ma w tym punkcie ekstremum lokalne, to

$$f'_x(x_0, y_0) = 0 \quad \text{i} \quad f'_y(x_0, y_0) = 0.$$

D o w ó d. Ponieważ funkcja f ma w punkcie (x_0, y_0) ekstremum, więc funkcja zmiennej x: $\varphi_1(x) = f(x, y_0)$ w punkcie x_0 oraz funkcja zmiennej y: $\varphi_2(y) = f(x_0, y)$ w punkcie y_0 mają również ekstrema (por. rys 6.2.). Zatem

$$\varphi_1'(x_0) = f'_x(x_0, y_0) = 0 \quad \text{oraz} \quad \varphi_2'(y_0) = f'_y(x_0, y_0) = 0,$$

co kończy dowód.



Rys 6.2.

Twierdzenie 6.1 orzeka, że dla funkcji mającej pochodne cząstkowe I rzędu, zerowanie się tych pochodnych w pewnym punkcie jest warunkiem koniecznym istnienia ekstremum lokalnego w tym punkcie.

Oczywiście funkcja może mieć ekstremum lokalne w punkcie, w którym nie ma pochodnych cząstkowych. Na przykład funkcja $z=\sqrt{x^2+y^2}$ ma w punkcie $(0,0)$ minimum, a jednocześnie w tym punkcie nie ma pochodnych cząstkowych.

Z twierdzenia 6.1. wynika jednak, że funkcja mająca pochodne cząstkowe na pewnym obszarze może mieć ekstrema lokalnie jedynie w tych punktach, w których obie pochodne cząstkowe są równe zeru. Punkty te nazywamy **punktami stacjonarnymi** funkcji.

PRZYKŁAD 6.1 Wykażemy, że funkcja

$$f(x,y)=y+x^2+\ln(y-x)$$

nie ma ekstremów lokalnych.

Funkcja f określona jest w każdym punkcie półpłaszczyzny $D=\{(x,y)\in R^2: y>x\}$ i w każdym punkcie tego obszaru ma pochodne cząstkowe. Zatem ekstrema lokalne tej funkcji mogą istnieć jedynie w punktach, w których obie pochodne cząstkowe są równe zeru. Obliczamy więc pochodne cząstkowe funkcji f i rozwiązujemy układ równań: $f'_x=0$, $f'_y=0$, czyli

$$\begin{cases} 2x-\dfrac{1}{y-x}=0, \\ 1+\dfrac{1}{y-x}=0. \end{cases}$$

Jedynym rozwiązaniem tego układu jest punkt $p_0=(-1/2,-3/2)$, który nie należy do dziedziny danej funkcji. Oznacza to, że funkcja f nie ma ekstremów lokalnych. ■

Zerowanie się pochodnych cząstkowych nie jest jednak warunkiem wystarczającym istnienia ekstremum, o czym przekonuje nas przykład funkcji $z(x,y)=x^2y$. Dla funkcji tej mamy: $z'_x(0,0)=0$, $z'_y(0,0)=0$ i $z(0,0)=0$. Jednocześnie łatwo widać, że na każdym otoczeniu punktu $(0,0)$ funkcja ta przyjmuje wartości dodatnie (gdy

$y>0$ i $x\neq 0$) i ujemne (gdy $y<0$ i $x\neq 0$), a zatem funkcja $z=x^2y$ nie ma ekstremum w punkcie $(0,0)$.

WARUNEK WYSTARCZAJĄCY ISTNIENIA EKSTREMUM. Załóżmy, że funkcja f zmiennych x i y ma ciągłe pochodne cząstkowe drugiego rzędu na pewnym obszarze D (krótko: f jest klasy C^2 na obszarze D). Funkcję W określoną wzorem

$$W(x,y)=f''_{xx}(x,y)f''_{yy}(x,y)-(f''_{xy}(x,y))^2,\quad (x,y)\in D,$$

nazywamy **wyróżnikiem funkcji** f.

TWIERDZENIE 6.2. Jeżeli funkcja f jest klasy C^2 na pewnym otoczeniu U punktu (x_0,y_0) oraz

1) $f'_x(x_0,y_0)=0$, $f'_y(x_0,y_0)=0$,

2) $W(x_0,y_0)>0$,

to funkcja f ma w punkcie (x_0,y_0) ekstremum lokalne właściwe, przy czym jest to maksimum, gdy $f''_{xx}(x_0,y_0)<0$ lub minimum, gdy $f''_{xx}(x_0,y_0)>0$.

Dowód. Z ciągłości pochodnych f''_{xx}, f''_{xy} i f''_{yy} na otoczeniu U wynika również ciągłość wyróżnika W na tym otoczeniu. Ponieważ $W(x_0,y_0)>0$, więc (zgodnie z twierdzeniem o lokalnym zachowaniu znaku) istnieje takie otoczenie $U_0\subset U$ punktu $p_0=(x_0,y_0)$, że $W(x,y)>0$ dla każdego $(x,y)\in U_0$. Ponadto zauważmy, że $f''_{xx}\neq 0$ dla $(x,y)\in U_0$ (gdyż w przeciwnym razie w punkcie, w którym $f''_{xx}=0$ mielibyśmy $W(x,y)\leq 0$).

Niech $p=(x_0+h,y_0+k)$, $h^2+k^2>0$, będzie dowolnym, różnym od p_0, punktem otoczenia U_0. Z twierdzenia Taylora (dla $n=2$) wynika, że istnieje liczba $\theta\in(0,1)$ taka, że

$$f(x_0+h,y_0+k)-f(x_0,y_0)=df(x_0,y_0)+\frac{1}{2}d^2f(x_0+\theta h,y_0+\theta k).$$

Ponieważ $f'_x(x_0,y_0)=0$ i $f'_y(x_0,y_0)=0$, więc

$$(1)\qquad f(x_0+h,y_0+k)-f(x_0,y_0)=\frac{1}{2}(f''_{xx}h^2+2f''_{xy}hk+f''_{yy}k^2)=$$
$$=\frac{1}{2f''_{xx}}((f''_{xx}h+f''_{xy}k)^2+k^2W),$$

przy czym występujące tu pochodne cząstkowe drugiego rzędu funkcji f, a także wyróżnik W obliczone są w punkcie $(x_0+\theta h, y_0+\theta k)$. Ponieważ $W>0$ oraz przynajmniej jeden z przyrostów h lub k jest różny od zera, więc

$$(2) \qquad (f''_{xx}h+f''_{xy}k)^2+k^2W>0.$$

Załóżmy, że $f''_{xx}(x_0,y_0)<0$. Ponieważ f''_{xx} jest ciągła i różna od zera na otoczeniu U_0, więc przyjmuje ona wartości ujemne w każdym punkcie tego otoczenia. Stąd

$$(3) \qquad f''_{xx}(x_0+\theta h, y_0+\theta k)<0.$$

Z (1), (2) i (3) otrzymujemy, że $f(x_0+h,y_0+k)-f(x_0,y_0)<0$. Zatem

$$\bigwedge_{p\in U(p_0)\setminus\{p_0\}} f(p)-f(p_0)<0,$$

a to oznacza, że funkcja f ma w punkcie $p_0=(x_0,y_0)$ maksimum lokalne właściwe.

Analogicznie, gdy $f''_{xx}(x_0,y_0)>0$ otrzymujemy, że funkcja f ma w punkcie $p_0=(x_0,y_0)$ minimum lokalne. □

Podobnie, korzystając ze wzoru Taylora, udowadniamy

TWIERDZENIE 6.3. Jeżeli funkcja f jest klasy C^2 na pewnym otoczeniu U punktu (x_0,y_0) oraz

1) $f'_x(x_0,y_0)=0, \quad f'_y(x_0,y_0)=0,$

2) $W(x_0,y_0)<0,$

to funkcja f w punkcie (x_0,y_0) nie ma ekstremum.

Natomiast w przypadku gdy w punkcie stacjonarnym funkcji wyróżnik jest równy zeru, funkcja może mieć, ale nie zawsze ma, ekstremum. Jest to tzw. przypadek wątpliwy, którym nie będziemy się zajmować. Rozstrzygnięcie, czy istnieje wtedy ekstremum funkcji, wymaga badania pochodnych cząstkowych wyższego rzędu lub korzystania bezpośrednio z definicji.

PRZYKŁAD 6.2. Sprawdzimy, w którym z punktów: $p_1=(-1,0)$, $p_2=(0,0)$, $p_3=(1,1)$ funkcja

$$z=x^7+2x^2y^5+y^6+x^5-16xy$$

ma ekstremum lokalne.

Rozważana funkcja jest klasy C^∞ na płaszczyźnie R^2. Obliczamy pochodne cząstkowe I i II rzędu:

$$z'_x=7x^6+4xy^5+5x^4-16y, \quad z'_y=10x^2y^4+6y^5-16x,$$

$$z''_{xx}=42x^5+4y^5+20x^3, \quad z''_{xy}=20xy^4-16, \quad z''_{yy}=40x^2y^3+30y^4.$$

Sprawdzamy najpierw, czy w punktach: p_1, p_2, p_3 spełniony jest układ równań: $z'_x=0$, $z'_y=0$, czyli

$$\begin{cases}7x^6+4xy^5+5x^4-16y=0,\\ 10x^2y^4+6y^5-16x=0.\end{cases}$$

Punkt $p_1=(-1,0)$ nie spełnia tego układu, zatem funkcja w punkcie p_1 nie ma ekstremum (nie jest spełniony warunek konieczny istnienia ekstremum). W punktach $p_2=(0,0)$, $p_3=(1,1)$ układ ten jest spełniony, zatem funkcja może (choć nie musi) mieć w tych punktach ekstrema lokalne. Obliczamy wartość wyróżnika $W(x,y)=z''_{xx}z''_{yy}-(z''_{xy})^2$ w punktach $p_2=(0,0)$ i $p_3=(1,1)$. Ponieważ

$$z''_{xx}(0,0)=0, \quad z''_{xy}(0,0)=-16, \quad z''_{yy}(0,0)=0,$$
$$z''_{xx}(1,1)=66, \quad z''_{xy}(1,1)=4, \quad z''_{yy}(1,1)=70,$$

więc

$$W(p_2)=W(0,0)=-256<0, \quad W(p_3)=W(1,1)=4604>0.$$

Stąd wynika, że w punkcie $p_2=(0,0)$ rozważana funkcja ekstremum nie ma. Natomiast w punkcie $p_3=(1,1)$ funkcja ta ma ekstremum lokalne i jest to minimum, gdyż $z''_{xx}(1,1)>0$. Obliczamy: $z_{min}=z(1,1)=-11$. ■

PRZYKŁAD 6.3. Wyznaczymy ekstrema lokalne funkcji określonej wzorem:

a) $z=(x^2-y^2)e^x$, b) $z=x^3-2x^2+y-\ln(y-x^2)$.

a) Funkcja ta jest funkcją klasy C^∞ (a więc również C^2) na zbiorze R^2, przy czym

$$z'_x=(2x+x^2-y^2)e^x, \quad z'_y=-2ye^x,$$

$$z''_{xx}=(2+4x+x^2-y^2)e^x, \quad z''_{xy}=-2ye^x, \quad z''_{yy}=-2e^x.$$

Wyznaczymy najpierw punkty stacjonarne funkcji , to znaczy rozwiążemy układ równań $z'_x = 0$, $z'_y = 0$, czyli

$$\begin{cases}(2x+x^2-y^2)e^x = 0,\\ -2ye^x = 0.\end{cases}$$

Układ ten ma dwa rozwiązania: $x=0$, $y=0$ oraz $x=-2$, $y=0$. Punkty $p_1=(0,0)$ i $p_2=(-2,0)$ są punktami stacjonarnymi badanej funkcji i jedynie w tych punktach funkcja ta może mieć ekstrema lokalne. Obliczamy dalej:

$$W(x,y)=z''_{xx}z''_{yy}-(z''_{xy})^2=-2(2+4x+x^2+y^2)e^{2x},$$

a następnie

$$W(p_1)=W(0,0)=-4,\quad W(p_2)=W(-2,0)=4e^{-4}.$$

Ponieważ $W(0,0)<0$, funkcja w punkcie $p_1=(0,0)$ ekstremum nie ma. Natomiast $W(-2,0)>0$, więc funkcja ta ma ekstremum lokalne w punkcie $p_2=(-2,0)$ i jest to maksimum (gdyż $z''_{xx}(-2,0)=-2e^{-2}<0$), przy czym $z_{max}=z(-2,0)=4e^{-2}$.

b) Funkcja ta jest funkcją klasy C^∞ (a więc również C^2) na zbiorze $D=\{(x,y)\in R^2:\ y>x^2\}$, przy czym

$$z'_x=3x^2-4x+\frac{2x}{y-x^2},\quad z'_y=1-\frac{1}{y-x^2},$$

$$z''_{xx}=6x-4+\frac{2y+2x^2}{(y-x^2)^2},\quad z''_{xy}=-\frac{2x}{(y-x^2)^2},\quad z''_{yy}=\frac{1}{(y-x^2)^2}.$$

Rozwiązujemy układ równań: $z'_x=0$, $z'_y=0$, czyli

$$\begin{cases}3x^2-4x+\dfrac{2x}{y-x^2}=0,\\ 1-\dfrac{1}{y-x^2}=0.\end{cases}$$

Układ ten dla $(x,y)\in D$ jest równoważny układowi:

$$\begin{cases}(3x^2-4x)(y-x^2)+2x=0,\\ y-x^2=1,\end{cases}$$

który ma dwa rozwiązania: $x=0$, $y=1$ oraz $x=2/3$, $y=13/9$. Łatwo sprawdzamy, że $p_1=(0,1)\in D$ i $p_2=(2/3,13/9)\in D$. Punkty p_1 i p_2 są punktami stacjonarnymi badanej funkcji i jedynie w tych punktach funkcja ta może mieć ekstrema lokalne. Następnie znajdujemy wartość wyróżnika

$$W(x,y)=z''_{xx}(x,y)z''_{yy}(x,y)-(z''_{xy}(x,y))^2$$

w punktach p_1 i p_2. Obliczamy kolejno:

$$z''_{xx}(p_1)=-2,\quad z''_{yy}(p_1)=1,\quad z''_{xy}(p_1)=0,$$
$$z''_{xx}(p_2)=34/9,\quad z''_{yy}(p_2)=1,\quad z''_{xy}(p_2)=-4/3,$$
$$W(p_1)=W(0,1)=-2,\quad W(p_2)=W(2/3,13/9)=2.$$

Ponieważ $W(0,1)<0$, więc dana funkcja w punkcie $p_1=(0,1)$ ekstremum nie ma . Natomiast $W(2/3,13/9)>0$, zatem funkcja ta ma ekstremum lokalne w punkcie $p_2=(2/3,13/9)$ i jest to minimum, gdyż $z''_{xx}(2/3,13/9)>0$. Na koniec obliczamy wartość badanej funkcji w punkcie p_2 i otrzymujemy $z_{min}=z(2/3,13/9)=23/27$ ■

WYZNACZANIE EKSTREMÓW ABSOLUTNYCH. Załóżmy, że dana jest funkcja dwóch zmiennych, ciągła na obszarze domkniętym i ograniczonym D. Z własności Weierstrassa funkcji ciągłych wynika, że funkcja ta osiąga na obszarze D swoją wartość najmniejszą i wartość największą (ekstrema absolutne). Aby wyznaczyć te wartości, znajdujemy najpierw takie punkty wewnętrzne obszaru D, w których funkcja ta może mieć ekstrema lokalne (tzn. punkty stacjonarne oraz punkty, w których funkcja nie ma pochodnych cząstkowych). Następnie rozważamy daną funkcję na brzegu obszaru D (mamy wtedy do czynienia z funkcją jednej zmiennej) i znajdujemy punkty, w których funkcja ta może osiągać wartość najmniejszą lub największą.

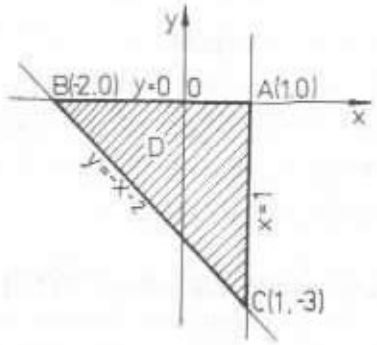
Wśród wyznaczonych w ten sposób punktów odnajdujemy już łatwo te, w których funkcja osiąga ekstrema absolutne na obszarze D.

PRZYKŁAD 6.4. Wyznaczymy ekstrema absolutne funkcji f na domkniętym obszarze D, gdy:

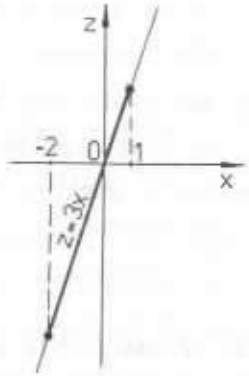
a) $f(x,y)=3x-3y+xy-y^2$, $D=\{(x,y)\in R^2:-2\le x\le 1\wedge -x-2\le y\le 0\}$.

b) $f(x,y)=xe^{-x^2-y^2}$, $D=\{(x,y)\in R^2: x^2+y^2-2x\le 0\}$.

a) Ponieważ funkcja f ma pochodne cząstkowe w dowolnym punkcie płaszczyzny R^2, więc funkcja ta może mieć ekstrema lokalne jedynie w punktach stacjonarnych. Rozwiązujemy układ równań: $f'_x=y+3=0$, $f'_y=x-2y-3=0$. Okazuje się, że punkt $(-3,-3)$ jest jedynym punktem stacjonarnym funkcji f, lecz nie należy on do wnętrza obszaru D. Tak więc, funkcja f nie ma ekstremów lokalnych wewnątrz obszaru D, a ekstrema absolutne osiąga na brzegu tego obszaru.



Rys 6.3.



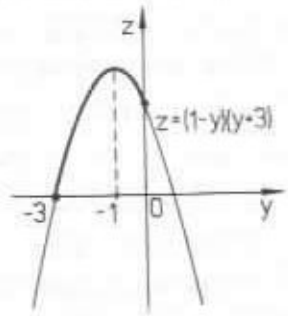
Rys 6.4.

Badamy funkcję f na kolejnych bokach trójkąta D (rys 6.3.) i wyznaczymy te punkty, w których funkcja przyjmuje wartości największe i najmniejsze na poszczególnych bokach. Spośród tych punktów wybierzemy potem te, w których osiągane są ekstrema absolutne funkcji f na obszarze D.

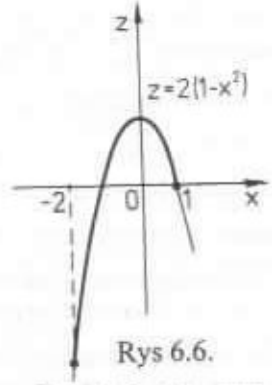
Na odcinku $\overline{AB}$ mamy: $y=0$, $-2\le x\le 1$, a funkcja f przyjmuje wartości $z(x)=3x$ (rys 6.4.). Funkcja ta osiąga swoją najmniejszą wartość dla $x=-2$, a największą dla $x=1$. Ekstremalne wartości funkcji f na odcinku $\overline{AB}$ są więc osiągane w punktach $p_1=(-2,0)$ i $p_2=(1,0)$.

Na odcinku $\overline{AC}$ mamy: $x=1$, $-3\le y\le 0$, a funkcja f ma wartości $z(y)=(1-y)(y+3)$. Jest to funkcja kwadratowa zmiennej y,

zatem łatwo widać (por. rys 6.5.), że wartość największą i najmniejszą funkcji $z(y)$ na przedziale $<-3,0>$ otrzymujemy odpowiednio dla $y=-1$ i $y=-3$. Funkcja f osiąga więc ekstrema na odcinku $\overline{AC}$ w punktach $p_3=(1,-1)$ i $p_4=(1,-3)$.



Rys 6.5.



Rys 6.6.

Na boku $\overline{BC}$ mamy: $y=-x-2$, $-2\le x\le 1$, zatem funkcja f przyjmuje wartości $z(x)=2(1-x^2)$. Funkcja $z(x)$ (por. rys 6.6.) osiąga na przedziale $<-2,1>$ największą wartość dla $x=0$, a najmniejszą dla $x=-2$. Zatem funkcja f na odcinku $\overline{BC}$ osiąga ekstrema w punktach $p_5=(0,-2)$ i $p_6=p_1=(-2,0)$.

Reasumując stwierdzamy, że funkcja $f(x,y)=3x-3y+xy-y^2$ może osiągać ekstrema absolutne na trójkącie D jedynie w punktach $p_1,\ldots,p_5$. Obliczamy wartości funkcji f w tych punktach:

$$f(p_1)=-6,\quad f(p_2)=3,\quad f(p_3)=4,\quad f(p_4)=0,\quad f(p_5)=2.$$

Stąd wynika, że maksimum absolutne na trójkącie D funkcja f osiąga w punkcie $p_3=(1,-1)$ i jest ono równe 4; minimum absolutne funkcji f na tym trójkącie jest równe -6 i osiągane jest w punkcie $p_1=(-2,0)$.

b) Funkcja ma pochodne cząstkowe w dowolnym punkcie płaszczyzny R^2, zatem ekstrema lokalne tej funkcji mogą istnieć jedynie w jej punktach stacjonarnych. Układ równań $f'_x=0$, $f'_y=0$, czyli

$$\begin{cases}(1-2x^2)e^{-x^2-y^2}=0,\\ -2xye^{-x^2-y^2}=0\end{cases}$$

ma dwa rozwiązania: $x=-1/\sqrt{2}$, $y=0$ oraz $x=1/\sqrt{2}$, $y=0$, przy czym tylko punkt $p_1=(1/\sqrt{2},0)$ jest punktem wewnętrznym obszaru D. W punkcie p_1 może istnieć ekstremum lokalne funkcji f i może to być ekstremum absolutne.

Obszar D jest kołem domkniętym o środku w punkcie $S=(1,0)$ i promieniu $r=1$. Brzegiem tego koła jest okrąg o równaniu $x^2+y^2-2x=0$. W punktach tego okręgu funkcja f przyjmuje wartości $z(x)=xe^{-2x}$, przy czym $x\in<0,2>$. Ponieważ $z'(x)=(1-2x)e^{-2x}=0$ dla $x=1/2$, więc funkcja $z(x)$ może osiągać swoją największą i najmniejszą wartość na przedziale $<0,2>$ w punkcie: $x=1/2$ (w tym punkcie może być ekstremum lokalne) lub w punktach końcowych przedziału: $x=0$, $x=2$. Zatem funkcja f może osiągać ekstrema absolutne na okręgu $x^2+y^2-2x=0$ w punktach

$$p_2=(1/2,\sqrt{3}/2),\quad p_3=(1/2,-\sqrt{3}/2),\quad p_4=(0,0),\quad p_5=(2,0).$$

Obliczamy wartości funkcji f w punktach $p_1,\dots,p_5$:

$$f(p_1)=1/\sqrt{2e},\quad f(p_2)=1/2e,\quad f(p_3)=1/2e,\quad f(p_4)=0,\quad f(p_5)=2/e^4.$$

Ponieważ

$$0<2/e^4<1/2e<1/\sqrt{2e}.$$

więc maksimum absolutne (największą wartość) na zbiorze D funkcja osiąga w punkcie $p_1=(1/\sqrt{2},0)$ i jest ono równe $1/\sqrt{2e}$, a minimum absolutne (najmniejszą wartość) na zbiorze D funkcja przyjmuje w punkcie $p_4=(0,0)$ i jest ono równe 0. ■

ZADANIA DO ROZWIĄZANIA.

1. Wykazać, że funkcja $z=z(x,y)$ nie ma ekstremów lokalnych:

a) $z=x^5+xy^2-y^3+2y$, b) $z=ye^{xy-x^2}$,

c) $z=y+\ln y-\frac{1}{y}\ln(1+x^2)$, d) $z=y^2-4y-\frac{x^2}{\sqrt{1-y}}$,

e) $z=x^2\ln(3+y^2)-y$, f) $z=\frac{x}{y}-x^2+\ln(4-y^2)$,

g) $z=\frac{2}{1-y}-\ln(1+x^2-y)$, h) $z=\frac{x-y}{e^{y-x^2}}$.

2. Sprawdzić, że funkcja $z=z(x,y)$ w punktach p_1 i p_2 nie ma ekstremów lokalnych, gdy:

a) $z=(x^2+y)e^{-\frac{1}{2}y^2}$, $p_1=(0,0)$, $p_2=(0,1)$,

b) $z=x^2y-x+y-ye^{xy^2}$, $p_1=(0,-1)$, $p_2=(1,0)$,

c) $z=x^4+y^4+xy^2+x^2y-2y^2-x$, $p_1=(0,0)$, $p_2=(0,-1)$,

d) $z=2x\sin y+y^2+\cos^2 x$, $p_1=(\pi,\pi)$, $p_2=(0,\pi)$.

3. Wykazać, że funkcja $z=z(x,y)$ ma w punkcie p_0 ekstremum lokalne i określić, czy jest to maksimum czy minimum, gdy:

a) $z=\frac{x}{y}+\frac{1}{x}+y$, $p_0=(1,1)$,

b) $z=y+2x^2-xy^2-x^2y-x^4-y^4$, $p_0=(1,0)$,

c) $z=x^7+xy^6+y^5+x^5-13x-11y$, $p_0=(1,1)$,

d) $z=ye^{-x^3-2y^2}$, $p_0=(0,1/2)$,

e) $z=y\cos^2 x-\cos y+x^2-y$, $p_0=(0,\pi)$.

4. Sprawdzić, w którym z punktów: p_1, p_2, p_3 funkcja $z=z(x,y)$ ma ekstremum lokalne, gdy:

a) $z=2x^3-2y^3+3x^2-3y^2$, $p_1=(0,0)$, $p_2=(0,-1)$, $p_3=(-1,0)$,

b) $z=xy+\frac{1}{x}+\frac{1}{y}$, $p_1=(-1,1)$, $p_2=(-1,-1)$, $p_3=(1,1)$,

c) $z=(2x+y^2)e^{-2x^2}$, $p_1=(-1/2,0)$, $p_2=(1/2,0)$, $p_3=(0,0)$.

5. Znaleźć, o ile istnieją, ekstrema lokalne funkcji określonej wzorem:

a) $z=x^2+2y^2+2xy-2x$, b) $z=x^4+2y^2-4xy+1$,

c) $z=6x^2y-4x^3-y^2-4y-3$, d) $z=x^3-y^3+3xy$,

e) $z = x^3 - y^3 + 3x^2 - 3xy + 3x - 3y$, f) $z = (y - x^2)e^{-y}$,

g) $z = 6y - 3x - 6xy + 3x^2 - x^3 - y^3$, h) $z = (1 - x^2 - y^2)e^{-y}$,

i) $z = \frac{y}{x} + \frac{1}{y} + x$, j) $z = 2xy + \frac{1}{x+y}$,

k) $z = \frac{1}{x+1} + \frac{1}{y} - xy - y$, l) $z = \frac{x^2 - y^2}{e^{y-2x}}$,

ł) $z = \ln(y + 2x) - 2x - y^2$, m) $z = xe^{-2x^2-y^2}$,

n) $z = \ln(4 - y - x^2) + y$, o) $z = \frac{2}{x} + \ln(x - y^2)$,

p) $z = 2y - \sin^2 y + (4x - x^4)\cos^2 y$, r) $z = (x^2 + y^2)^2 + xy$.

6. Znaleźć najmniejszą i największą wartość funkcji $z = z(x,y)$ na obszarze domkniętym D, gdy:

a) $z = 2x^2 - y^2 - 6x$, $D = \{(x,y) \in R^2: x^2 + y^2 \le 9 \wedge x \ge 0\}$,

b) $z = 2 + x - y - xy - y^2$, $D = \{(x,y) \in R^2: 0 \le x \le 1 \wedge 0 \le y \le 1 - x\}$

c) $z = 2x^3 - y^3 - 6x + 3y$, $D = \{(x,y) \in R^2: -2 \le x \le 0 \wedge 0 \le y \le 2\}$

d) $z = x^3 + 6xy + y^3$, $D = \{(x,y) \in R^2: -1 \le x \le 1 \wedge -1 \le y \le 2\}$,

e) $z = 3x^2 - y^2$, $D = \{(x,y) \in R^2: x^2 + y^2 \le 4\}$,

f) $z = 2x^3 - y^3 - 3x^2y + 6y$, $D = \{(x,y) \in R^2: -1 \le x \le 1 \wedge 0 \le y \le 2\}$,

g) $z = xy - \frac{1}{x} - \frac{1}{y}$, $D = \{(x,y) \in R^2: 1 \le x \le 2 \wedge 1 \le y \le 2\}$,

h) $z = ye^{x^2+y^2-2y}$, $D = \{(x,y) \in R^2: x^2 + y^2 \le 1\}$,

i) $z = 2\ln(x + y + 2) - x + y^2$, $D = \{(x,y) \in R^2: 0 \le y \le 1 \wedge y - 1 \le x \le 1 - y\}$,

j) $z = xe^{-2x^2-y^2}$, $D = \{(x,y) \in R^2: x^2 + y^2 \le 1\} \wedge y \ge 0\}$.

Odpowiedzi.

1. Wsk. Wystarczy wykazać, że funkcja nie ma punktów stacjonarnych (tzn., że układ $z'_x = 0$, $z'_y = 0$ nie ma rozwiązań w dziedzinie danej funkcji) albo, gdy funkcja ma punkt stacjonarny p_0, to $W(p_0) < 0$.

2. Wsk. Wystarczy sprawdzić, że podany punkt p_k nie jest punktem stacjonarnym, tzn. $z'_x(p_k) \neq 0$ lub $z'_y(p_k) \neq 0$ albo, gdy podany punkt p_k jest punktem stacjonarnym funkcji, że $W(p_k) < 0$.

3. a) min. lok., b) max. lok., c) min. lok. d) max. lok., e) max. lok.

4. a) min. lok. w punkcie p_2, max. lok. w punkcie p_3,
b) min. lok. w punkcie p_2, c) min. lok. w punkcie p_1.

5. a) $z_{min} = z(2,-1) = -2$, b) $z_{min} = z(1,1) = z(-1,-1) = 0$, c) $z_{max} = z(0,-2) = 1$,
d) $z_{min} = z(1,-1) = -1$, e) $z_{max} = z(-2,1) = 0$, f) $z_{max} = z(0,1) = 1/e$,
g) $z_{min} = z(-1,-2) = -9$, h) $z_{max} = z(0,1-\sqrt{2}) = (2\sqrt{2} - 2)e^{\sqrt{2}-1}$, i) $z_{min} = z(1,1) = 3$,
j) nie ma, k) $z_{max} = z(-2,-1) = -3$, l) $z_{max} = z(-4/3,-2/3) = 4/3e^2$,
l) $z_{max} = z(1/4,1/2) = -3/4$, m) $z_{max} = z(1/2,0) = 1/2\sqrt{e}$,
$z_{min} = z(-1/2,0) = -1/2\sqrt{e}$, n) $z_{max} = z(0,3) = 3$, o) nie ma,
p) $z_{max} = z(1, \frac{\pi}{12} + k\pi) = 1 + \sqrt{3} + \frac{\pi}{6} + 2k\pi$, $k \in C$,
r) $z_{min} = z(1/2\sqrt{2}, -1/2\sqrt{2}) = z(-1/2\sqrt{2}, 1/2\sqrt{2}) = -1/16$.

6. a) wartość najmniejsza: $z(1,\sqrt{8}) = z(1,-\sqrt{8}) = -12$,
wartość największa: $z(0,0) = z(3,0) = 0$,
b) wartość najmniejsza: $z(0,1) = 0$, wartość największa: $z(1,0) = 3$,
c) wartość najmniejsza: $z(-2,2) = -6$, wartość największa: $z(-1,1) = 6$,
d) wartość najmniejsza: $z(-1,\sqrt{2}) = -1 - 4\sqrt{2}$, wartość największa: $z(1,2) = 21$,
e) wartość najmniejsza: $z(0,-2) = z(0,2) = -4$,
wartość największa: $z(-2,0) = z(2,0) = 12$,
f) wartość najmniejsza: $z(-1,2) = -4$, wartość największa: $z(0,\sqrt{2}) = 4\sqrt{2}$,
g) wartość najmniejsza: $z(1,1) = -1$, wartość największa: $z(2,2) = 3$,
h) wartość najmniejsza: $z(0,-1) = -e^3$, wartość największa: $z(-\sqrt{3}/2,1/2) =$
$= z(\sqrt{3}/2,1/2) = 1/2$,
i) wartość najmniejsza: $z(-1,0) = 1$, wartość największa: $z(0,1) = 1 + \ln 9$,
j) wartość najmniejsza: $z(-1/2,0) = -1/2\sqrt{e}$, wartość największa: $z(1/2,0) = 1/2\sqrt{e}$.

7. FUNKCJA UWIKŁANA.

FUNKCJA UWIKŁANA. Weźmy pod uwagę równania:

$$2x-y+1=0, \quad 2x-y^2+4=0, \quad x^2+y^2+1=0.$$

Pierwsze z nich określa na zbiorze R dokładnie jedną funkcję $y=y(x)$, dla której $2x-y(x)+1=0$ dla każdego $x\in R$. Jest to funkcja postaci: $y=2x+1$. Drugie równanie na przedziale $I=<-2,+\infty)$ określa dokładnie dwie funkcje c i ą g ł e $y=y(x)$ takie, że $2x-(y(x))^2+4=0$ dla $x\in I$. Inaczej powiemy: istnieją dokładnie dwie funkcje ciągłe na przedziale I spełniające dane równanie. Są to funkcje: $y=\sqrt{2x+4}$ i $y=-\sqrt{2x+4}$. Trzecie z równań nie określa żadnej funkcji, gdyż dla każdego $(x,y)\in R^2$ mamy $x^2+y^2+1\neq 0$. Każdą z funkcji wyżej określonych nazywamy funkcją uwikłaną wyznaczoną przez odpowiednie równanie. Ostatnie równanie nie określa żadnej funkcji uwikłanej.

Zauważmy jeszcze, że drugie z rozważanych równań jest spełnione przez nieskończenie wiele funkcji nieciągłych. Na przykład każda funkcja postaci:

$$y(x)=\begin{cases}\sqrt{2x+4}, & x\in<-2,a),\\ -\sqrt{2x+4}, & x\in<a,+\infty),\end{cases}$$

gdzie a jest dowolną liczbą z przedziału $(-2,+\infty)$, spełnia to równanie. W dalszym ciągu ograniczymy się do rozważania tylko funkcji ciągłych wyznaczonych przez odpowiednie równania.

Niech F będzie funkcją dwóch zmiennych określoną na pewnym obszarze. Każdą funkcję $y=y(x)$ c i ą g ł ą na pewnym przedziale I taką, że

$$\bigwedge_{x\in I} F(x,y(x))=0$$

nazywamy **funkcją uwikłaną** określoną równaniem

$$F(x,y)=0. \tag{7.1}$$

Na podstawie podanych wyżej przykładów stwierdzamy, że równanie $F(x,y)=0$ może określać na pewnym przedziale jedną lub więcej funkcji uwikłanych, albo też może nie określać żadnej funkcji uwikłanej.

Równanie (7.1) nazywane jest też p o s t a c i ą u w i k ł a n ą funkcji $y=y(x)$.

Załóżmy teraz, że $F(x_0,y_0)=0$, czyli punkt (x_0,y_0) spełnia równanie (7.1). Postawmy pytanie: kiedy równanie $F(x,y)=0$ wyznacza dokładnie jedną funkcję uwikłaną $y=y(x)$ taką, że $y(x_0)=y_0$, czyli taką funkcję uwikłaną, której wykres przechodzi przez punkt (x_0,y_0). Na pytanie to odpowiada następujące twierdzenie.

TWIERDZENIE 7.1 (o istnieniu funkcji uwikłanej). Jeżeli F jest funkcją klasy C^1 na pewnym otoczeniu punktu (x_0,y_0) oraz

$$F(x_0,y_0)=0 \quad \text{i} \quad F'_y(x_0,y_0)\neq 0,$$

to na pewnym otoczeniu punktu x_0 istnieje dokładnie jedna funkcja uwikłana $y=y(x)$ określona równaniem $F(x,y)=0$ spełniająca warunek $y(x_0)=y_0$, przy czym funkcja ta ma ciągłą pochodną określoną wzorem

$$y'(x)=-\frac{F'_x(x,y(x))}{F'_y(x,y(x))}. \tag{7.2}$$

Dowód tego twierdzenia pomijamy. Zauważmy jedynie, że wzór na pochodną funkcji uwikłanej otrzymujemy łatwo różniczkując równość $F(x,y(x))=0$. Mamy wówczas

$$F'_x+F'_y\,y'=0, \tag{7.3}$$

a stąd wynika natychmiast wzór (7.2).

P R Z Y K Ł A D 7.1. Weźmy pod uwagę równanie

$$x^2y^2-2e^{x+y}+x=0. \tag{1}$$

Funkcja $F(x,y)=x^2y^2-2e^{x+y}+x$ jest funkcją ciągłą wraz ze swoimi pochodnymi cząstkowymi na całej płaszczyźnie. Ponadto $F(1,-1)=0$ oraz $F'_y(1,-1)=-4\neq 0$. Zgodnie z twierdzeniem 7.1 na pewnym otoczeniu punktu $x_0=1$ istnieje dokładnie jedna funkcja uwikłana $y=y(x)$ określona równaniem (1), dla której $y(1)=-1$. Z twierdzenia 7.1 wynika również, że

$$y'(x)=-\frac{2xy^2-2e^{x+y}+1}{2x^2y-2e^{x+y}},$$

gdzie $y = y(x)$. Stąd otrzymujemy $y'(1) = 1/4$. W konsekwencji, wobec ciągłości $y'(x)$, mamy, że $y'(x) > 0$ na pewnym otoczeniu punktu $x_0 = 1$. Oznacza to, że funkcja uwikłana, o której mówimy, jest funkcją rosnącą na tym otoczeniu. ■

Przy dodatkowym założeniu, że funkcja F w równaniu (7.1) jest klasy C^2, otrzymujemy wzór na drugą pochodną funkcji uwikłanej. W tym celu różniczkujemy względem x równość (7.3). Wówczas mamy

$$F''_{xx} + F''_{xy}y' + (F''_{yx} + F''_{yy}y')y' + F'_y y'' = 0,$$

a stąd

$$y''(x) = -\frac{F''_{xx} + 2F''_{xy}y' + F''_{yy}(y')^2}{F'_y}.$$

Uwzględniając w powyższym wzór (7.2) mamy

$$y''(x) = -\frac{F''_{xx}(F'_y)^2 - 2F''_{xy}F'_xF'_y + F''_{yy}(F'_x)^2}{(F'_y)^3},$$

przy czym F'_x, F'_y, F''_{xx}, F''_{xy}, F''_{yy} oznaczają tu pochodne cząstkowe w punkcie $(x, y(x))$.

PRZYKŁAD 7.2. Łatwo sprawdzić, że równanie

$$(1) \qquad 2x + y - \cos xy - 1 = 0$$

określa na pewnym otoczeniu punktu $x_0 = 1$ dokładnie jedną funkcję uwikłaną $y = y(x)$ spełniającą warunek $y(1) = 0$. Istotnie, dla funkcji $F(x,y) = 2x + y - \cos xy - 1$ mamy $F(1,0) = 0$ i $F'_y(1,0) = 1 \neq 0$.

Ponieważ

$$F'_x = 2 + y\sin xy, \quad F'_y = 1 + x\sin xy,$$

$$F''_{xx} = y^2\cos xy, \quad F''_{xy} = \sin xy + xy\cos xy, \quad F''_{yy} = x^2\cos xy$$

oraz

$$F'_x(1,0) = 2, \; F'_y(1,0) = 1, \; F''_{xx}(1,0) = 0, \; F''_{xy}(1,0) = 0, \; F''_{yy}(1,0) = 1,$$

więc $y'(1) = -2$, $y''(1) = -4$.

Stąd, wobec ciągłości y' i y'', wynika, że rozważana funkcja uwikłana $y = y(x)$ na pewnym otoczeniu punktu $x_0 = 1$ jest malejąca, a jej wykres jest krzywą wklęsłą. ■

EKSTREMA FUNKCJI UWIKŁANEJ. Jak pokazaliśmy w przykładzie 7.2, nawet w przypadku gdy nie potrafimy efektywnie rozwiązać równania $F(x,y) = 0$ względem zmiennej y, możemy podać pewne informacje o funkcji uwikłanej $y = y(x)$ określonej tym równaniem. Teraz pokażemy w jaki sposób można wyznaczyć ekstrema lokalne funkcji uwikłanej nie znając jej postaci.

Z warunku wystarczającego istnienia ekstremum lokalnego funkcji jednej zmiennej (tom 1, rozdz. III, 5) i twierdzenia 7.1 o istnieniu funkcji uwikłanej łatwo wynika następujące twierdzenie.

TWIERDZENIE 7.2. (warunek wystarczający istnienia ekstremum funkcji uwikłanej). Jeżeli F jest funkcją klasy C^2 na pewnym otoczeniu punktu (x_0, y_0) oraz

1) $F(x_0,y_0) = 0$, $F'_x(x_0,y_0) = 0$, $F'_y(x_0,y_0) \neq 0$,

2) $I(x_0,y_0) = -\dfrac{F''_{xx}(x_0,y_0)}{F'_y(x_0,y_0)} \neq 0$,

to funkcja uwikłana $y = y(x)$ określona równaniem $F(x,y) = 0$ i spełniająca warunek $y(x_0) = y_0$, ma w punkcie x_0 ekstremum lokalne równe y_0, przy czym jest to maksimum, gdy $I(x_0,y_0) < 0$ lub minimum, gdy $I(x_0,y_0) > 0$.

Dowód. Z założenia 1) i twierdzenia 7.1. wynika istnienie funkcji uwikłanej $y = y(x)$ takiej, że $y(x_0) = y_0$, a także zerowanie się pierwszej pochodnej tej funkcji w punkcie x_0. Uwzględniając, że $F'_x(x_0,y_0) = 0$ otrzymujemy

$$y''(x_0) = -\frac{F''_{xx}(x_0,y_0)}{F'_y(x_0,y_0)}.$$

Z założenia 2) wynika, że $y''(x_0) \neq 0$, a zatem funkcja $y(x)$ ma w punkcie x_0 ekstremum lokalne. Jest to maksimum, gdy $y''(x_0) = I(x_0,y_0) < 0$ lub minimum, gdy $y''(x_0) = I(x_0,y_0) > 0$, co kończy dowód. □

PRZYKŁAD 7.3. Wyznaczymy ekstrema lokalne funkcji $y = y(x)$ określonych równaniem:

a) $2x^4 + x^2y - y^2 + 3y - 2 = 0$,

b) $e^{x-1} + e^y - x + y - 1 = 0$.

a) Przy oznaczeniu $F(x,y)=2x^4+x^2y-y^2+3y-2=0$, mamy

$$F'_x=8x^3+2xy,\quad F'_y=x^2-2y+3,\quad F''_{xx}=24x^2+2y$$

oraz

$$I(x,y)=-\frac{24x^2+2y}{x^2-2y+3}.$$

Najpierw rozwiązujemy układ równań $F(x,y)=0$, $F'_x(x,y)=0$:

$$\begin{cases}2x^4+x^2y-y^2+3y-2=0,\\ 8x^3+2xy=0.\end{cases}$$

Układ ten jest równoważny alternatywie:

$$(1)\qquad \begin{cases}2x^4+x^2y-y^2+3y-2=0,\\ y=-4x^2,\end{cases}$$

lub

$$(2)\qquad \begin{cases}2x^4+x^2y-y^2+3y-2=0,\\ x=0.\end{cases}$$

Nietrudno sprawdzić, że układ (1) jest sprzeczny, natomiast układ (2) ma dwa rozwiązania $x=0$, $y=1$ oraz $x=0$, $y=2$. Sprawdzamy, że $F'_y(0,1)=1\neq 0$ i $F'_y(0,2)=-1\neq 0$. Ponieważ

$$I(0,1)=-2<0,\qquad I(0,2)=4>0,$$

zatem funkcja uwikłana $y=y(x)$ określona na pewnym otoczeniu punktu $x_0=0$, dla której $y(0)=1$, ma w punkcie $x_0=0$ maksimum lokalne równe 1, zaś funkcja uwikłana $y=y(x)$ taka, że $y(0)=2$ ma w punkcie $x_0=0$ minimum lokalne równe 2.

Wynik ten oznacza, że podane równanie określa na pewnym otoczeniu punktu $x_0=0$ dwie funkcje uwikłane $y=y(x)$. Jedna z nich spełnia warunek $y(0)=1$, a druga $y(0)=2$. Pierwsza z nich ma w punkcie $x_0=0$ maksimum lokalne, a druga ma w tym punkcie minimum lokalne.

Rozważane równanie jest równaniem kwadratowym ze względu na zmienną y i dlatego nietrudno jest znaleźć te dwie funkcje. Są to:

$y=1-x^2$, $x\in R$ oraz $y=2+2x^2$, $x\in R$. Czytelnikowi proponujemy naszkicowanie wykresów tych funkcji.

b) Dla $F(x,y)=e^{x-1}+e^y-x+y-1$ mamy

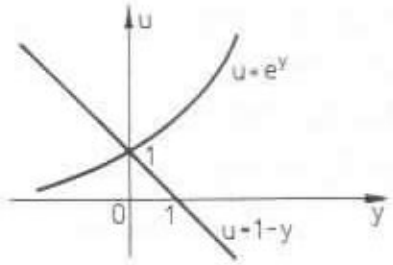
$$F'_x=e^{x-1}-1,\quad F'_y=e^y+1,\quad F''_{xx}=e^{x-1},\quad I(x,y)=-\frac{e^{x-1}}{e^y+1}.$$

Rozwiązujemy układ równań $F(x,y)=0$, $F'_x(x,y)=0$, czyli

$$(1)\qquad \begin{cases}e^{x-1}+e^y-x+y-1=0,\\ e^{x-1}-1=0.\end{cases}$$

Ponieważ drugie równanie układu jest spełnione jedynie dla $x=1$, więc (wstawiając $x=1$ do pierwszego równania) otrzymujemy układ równoważny postaci

$$\begin{cases}e^y=1-y,\\ x=1.\end{cases}$$



Rys 7.1.

Pierwsze równanie tego układu jest spełnione jedynie (por. rys 7.1.) dla $y=0$. Stąd wynika, że układ równań (1) ma jedno rozwiązanie: $x=1$, $y=0$. Sprawdzamy, że $F'_y(1,0)=2\neq 0$. Ponieważ $I(1,0)=-1/2<0$, więc funkcja uwikłana $y=y(x)$ określona rozważanym równaniem i przechodząca przez punkt (1,0) ma w punkcie $x_0=1$ maksimum lokalne równe 0. ■

Podobnie funkcję ciągłą $z=z(x,y)$ spełniającą równość $F(x,y,z(x,y))=0$ w każdym punkcie pewnego obszaru nazywamy funkcją uwikłaną dwóch zmiennych określoną równaniem $F(x,y,z)=0$. Na przykład równanie $xy+2z-x^2=0$ na obszarze R^2 określa jedną funkcję uwikłaną $z=z(x,y)$ i jest ona postaci: $z=(x^2-xy)/2$, natomiast

równanie $x^2+y^2+z^2-4=0$ określa na kole $K=\{(x,y)\in R^2: x^2+y^2\le 4\}$ dwie funkcje uwikłane $z=z(x,y)$ postaci: $z=\sqrt{4-x^2-y^2}$, $z=-\sqrt{4-x^2-y^2}$. Równanie $x^2+y^2+z^2+1=0$ nie określa żadnej funkcji uwikłanej.

Zainteresowanych zagadnieniem funkcji uwikłanych wielu zmiennych odsyłamy do obszerniejszych podręczników.

ZADANIA DO ROZWIĄZANIA.

1. Znaleźć pochodną $y'(x)$ funkcji uwikłanych określonych równaniami:
   a) $x^2+y^3+xy^5+1=0$, b) $x^2y+\ln y-2x=0$,
   c) $x^4+y^4-x^2-y^2=0$, d) $x-y+\ln(x^2+y^2)=1$,
   e) $xe^{x-y}-y=0$, f) $x\sin(x-2y)-y=1$.
2. Sprawdzić, że podane równanie określa funkcję uwikłaną $y=y(x)$ taką, że $y(x_0)=y_0$, a następnie obliczyć $y'(x_0)$ i $y''(x_0)$:
   a) $x-e^{2y}+ye^{x-1}=0$, $y(1)=0$,
   b) $e^{2y-x}-y^3+x=2$, $y(2)=1$,
   c) $xy+\cos x-\sin y=1$, $y(0)=0$,
   d) $e^xy^2-xe^{xy}-y=0$, $y(0)=1$,
   e) $x^2\ln(1+y^2)-y^3+\ln x+2y=0$, $y(1)=0$.
3. Sprawdzić, że równanie $y+x^2-y\cos^2 x+\sin y=0$ określa na pewnym otoczeniu punktu $x_0=0$ dokładnie jedną funkcję uwikłaną $y=y(x)$, dla której $y(0)=-\pi$. Wykazać, że funkcja ta osiąga w punkcie $x_0=0$ maksimum lokalne.
4. Sprawdzić, że równanie $ye^{xy}-x^2e^y+x^2-xy^3-1=0$ określa na pewnym otoczeniu punktu $x_0=0$ dokładnie jedną funkcję uwikłaną $y=y(x)$ taką, że $y(0)=1$. Wykazać, że funkcja ta ma w punkcie $x_0=0$ minimum lokalne.
5. Wykazać, że funkcja $y=y(x)$ dana równaniem
$$x\ln(1+y^2)+y^3\ln(2x-x^2)+y+x^3-x=0$$

i spełniająca warunek $y(1)=0$ jest na pewnym otoczeniu punktu $x_0=1$ funkcją malejąca, a jej wykres jest krzywą wklęsłą.

6. Sprawdzić, że przez punkt (1,1) przechodzi dokładnie jedna funkcja uwikłana $y=y(x)$ dana równaniem
$$y^2\ln(2-x^2)-x\ln y+x^3-y=0.$$
Wykazać, że na pewnym otoczeniu punktu $x_0=1$ funkcja ta jest rosnąca, a jej wykres jest krzywą wklęsłą.
7. Sprawdzić, że równanie $x^2e^{xy}+ye^x+x+y^2=0$ określa na pewnym otoczeniu punktu $y_0=0$ dokładnie jedną funkcję uwikłaną $x=x(y)$ taką, że $x(0)=-1$, a następnie obliczyć $x'(0)$ i $x''(0)$.
8. Wyznaczyć ekstrema lokalne funkcji uwikłanych $y=y(x)$ określonych równaniami:
   a) $x^2+y^2-2x-2y+1=0$, b) $e^{xy}-xy+2y-3=0$,
   c) $x^2-2xy-3y^2-2x-6y+1=0$, d) $x^3+y^3+3xy=0$,
   e) $x^4+y^2=4xy$, f) $x^2+y^2+(x+y)^2=6$,
   g) $2x^4+3x^2y-2x^2+y^2-y=0$, h) $x^4-2x^2-y^2-4xy+1=0$,
   i) $x^4-x^2-y^2-y=0$, j) $x^4-2x^2y-x^2+y^2+y=0$.

Odpowiedzi.

1. a) $y'=-\dfrac{2x+y^5}{3y^2+5xy^4}$, b) $y'=-\dfrac{2y(xy-1)}{x^2y+1}$, c) $y'=\dfrac{x(2x^2-1)}{y(2y^2-1)}$,
   d) $y'=\dfrac{x^2+y^2+2x}{x^2+y^2-2y}$, e) $y'=\dfrac{1+x}{x+e^{x-y}}$, f) $y'=\dfrac{\sin(x-2y)+x\cos(x-2y)}{1+2x\cos(x-2y)}$.
2. a) $y'(1)=1$, $y''(1)=-2$; b) $y'(2)=0$, $y''(2)=1$; c) $y'(0)=0$, $y''(0)=-1$,
   d) $y'(0)=0$, $y''(0)=1$; e) $y'(1)=-1/2$, $y''(1)=1/4$.
8. a) $y_{max}(1)=2$, $y_{min}(1)=0$; b) $y_{max}(0)=1$; c) $y_{min}(1)=0$, $y_{max}(-1)=-2$;
   d) $y_{min}(-\sqrt[3]{2})=-\sqrt[3]{4}$; e) $y_{max}(\sqrt{3})=3\sqrt{3}$, $y_{min}(-\sqrt{3})=-3\sqrt{3}$; f) $y_{min}(1)=-2$, $y_{max}(-1)=2$; g) $y_{max}(0)=0$, $y_{max}(0)=1$; h) $y_{min}(1)=0$, $y_{max}(-1)=0$;
   i) $y_{max}(0)=0$, $y_{min}(0)=-1$; j) $y_{min}(0)=0$, $y_{min}(0)=-1$

III. RACHUNEK CAŁKOWY FUNKCJI WIELU ZMIENNYCH

1. OKREŚLENIE CAŁKI PODWÓJNEJ I JEJ INTERPRETACJA GEOMETRYCZNA.

OBSZAR NORMALNY WZGLĘDEM OSI. OBSZAR REGULARNY. Definicje obszaru i obszaru domkniętego podane w rozdziale II uzupełnimy teraz definicjami obszaru normalnego względem osi 0x (0y) i obszaru regularnego.

Obszar domknięty $D \subset R^2$ nazywamy **normalnym względem osi 0x**, jeśli

$$D = \{(x,y) \in R^2 : a \le x \le b \wedge g(x) \le y \le h(x)\},$$

gdzie g i h są funkcjami ciągłymi dla $x \in \langle a,b \rangle$ i $g(x) < h(x)$ dla $x \in (a,b)$.

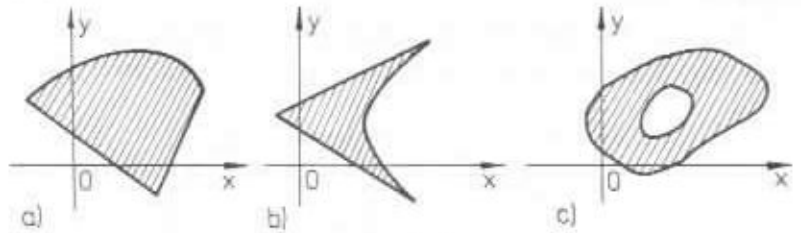
Obszar domknięty $D \subset R^2$ nazywamy **normalnym względem osi 0y**, jeśli

$$D = \{(x,y) \in R^2 : c \le y \le d \wedge k(y) \le x \le l(y)\},$$

gdzie k i l są funkcjami ciągłymi dla $y \in \langle c,d \rangle$ i $k(y) < l(y)$ dla $y \in (c,d)$.

Obszar domknięty $D \subset R^2$ jest normalny względem osi 0x (0y) wtedy i tylko wtedy, gdy prosta prostopadła do osi 0x (0y) i przechodząca przez punkt wewnętrzny obszaru D przecina brzeg tego obszaru dokładnie w dwóch punktach.

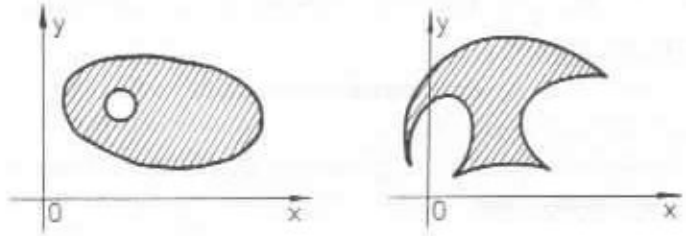
Na rysunku 1.1 podane są przykłady a) obszaru normalnego względem obu osi, b) obszaru normalnego tylko względem osi 0y, c) obszaru, który nie jest normalny ani względem osi 0x, ani też względem osi 0y.



Rys 1.1

Domknięty obszar $D \subset R^2$ nazywać będziemy obszarem **regularnym**, jeżeli można go przedstawić jako sumę skończonej liczby obszarów normalnych względem osi 0x lub 0y i takich, że żadne dwa z tych obszarów nie mają wspólnych punktów wewnętrznych.

Dowolny obszar normalny względem przynajmniej jednej z osi układu jest oczywiście obszarem regularnym. Na rysunku 1.2 przedstawione są inne przykłady obszarów regularnych.



Rys 1.2

Z przyjętych definicji wynika, że obszar regularny jest obszarem domkniętym i ograniczonym.

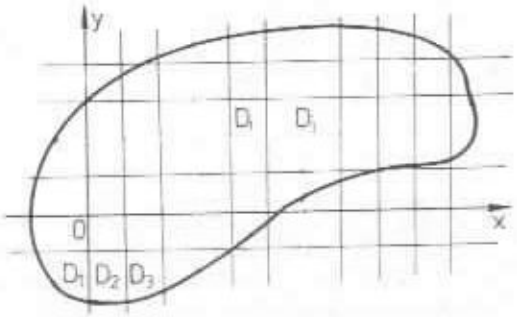
OKREŚLENIE CAŁKI PODWÓJNEJ. Niech f będzie funkcją określoną na regularnym obszarze $D \subset R^2$. Podzielmy obszar D w dowolny sposób (np. prostymi równoległymi do osi 0x i 0y) na n domknię-

tych obszarów częściowych D_i odpowiednio o polach ΔD_i, $i=1,2,\ldots,n$, w ten sposób, by

1° żadne dwa obszary D_i, D_j, $i \neq j$ nie miały wspólnych punktów wewnętrznych,

2° suma obszarów częściowych D_i była obszarem D (rys 1.3)

$$D = D_1 \cup D_2 \cup \cdots \cup D_n .$$



Rys 1.3

Niech d_i oznacza kres górny odległości dwu dowolnych punktów obszaru D_i; jeśli obszar D_i jest prostokątem o długościach boków $\Delta x_i, \Delta y_i$, to d_i jest długością przekątnej tego prostokąta, a pole $\Delta D_i = \Delta x_i \, \Delta y_i$. Liczbę

$$\delta_n = \max(d_1, d_2, \ldots, d_n)$$

nazywa się średnicą tego podziału.

W każdym obszarze częściowym D_i obierzmy punkt pośredni $(\bar{x}_i, \bar{y}_i)$ i utwórzmy sumę

$$S_n = f(\bar{x}_1, \bar{y}_1)\Delta D_1 + f(\bar{x}_2, \bar{y}_2)\Delta D_2 + \cdots + f(\bar{x}_n, \bar{y}_n)\Delta D_n = \sum_{i=1}^{n} f(\bar{x}_i, \bar{y}_i)\Delta D_i .$$

Sumę S_n nazywamy sumą całkową (sumą całkową Riemanna) funkcji f na obszarze D.

Utwórzmy teraz normalny ciąg podziałów obszaru D na obszary częściowe, to znaczy taki ciąg, że $\delta_n \to 0$ przy $n \to \infty$; ciągowi temu odpowiada ciąg (S_n) sum całkowych funkcji f. Rozważmy granicę

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{i=1}^{n} f(\bar{x}_i, \bar{y}_i)\Delta D_i .$$

Jeżeli dla każdego normalnego ciągu podziałów obszaru D i każdego wyboru punktów pośrednich $(\bar{x}_i, \bar{y}_i)$ w obszarach częściowych istnieje ta sama skończona granica ciągu (S_n), to tę granicę nazywamy **całką podwójną** funkcji f na obszarze D i oznaczamy symbolem

$$\iint_D f(x,y)\,dxdy .$$

Zatem

$$\iint_D f(x,y)\,dxdy \overset{\text{def}}{=} \lim_{n\to\infty} \sum_{i=1}^{n} f(\bar{x}_i, \bar{y}_i)\Delta D_i .$$

Funkcję f, dla której istnieje całka podwójna na obszarze D, nazywamy funkcją całkowalną (w sensie Riemanna) na obszarze D.

Całkę podwójną funkcji f na zbiorze D, który jest sumą n regularnych obszarów $D_1, \ldots, D_n$, przy czym żadne dwa z tych obszarów nie mają wspólnych punktów wewnętrznych, określamy równością:

$$\iint_D f(x,y)\,dxdy \overset{\text{def}}{=} \sum_{i=1}^{n} \iint_{D_i} f(x,y)\,dxdy .$$

TWIERDZENIE 1.1 (warunek konieczny całkowalności). Jeżeli funkcja f jest całkowalna na regularnym obszarze D, to jest funkcją ograniczoną na tym obszarze.

Stąd wynika, że funkcja nieograniczona na obszarze D nie jest całkowalna na tym obszarze. Można jednak w tym przypadku mówić o całce podwójnej *niewłaściwej*. W tej książce temat ten pomijamy.

TWIERDZENIE 1.2 (warunki wystarczające całkowalności).

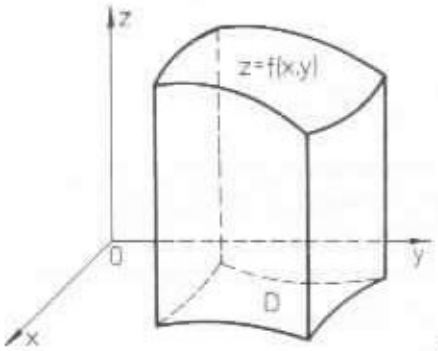
(1) Jeżeli funkcja f jest ciągła na regularnym obszarze D, to jest całkowalna na tym obszarze.

(2) Jeżeli funkcja f jest ograniczona na regularnym obszarze D i jest ciągła na tym obszarze z wyjątkiem punktów leżących na skończonej liczbie wykresów funkcji postaci $y = y(x)$ lub $x = x(y)$, to f jest funkcją całkowalną na obszarze D.

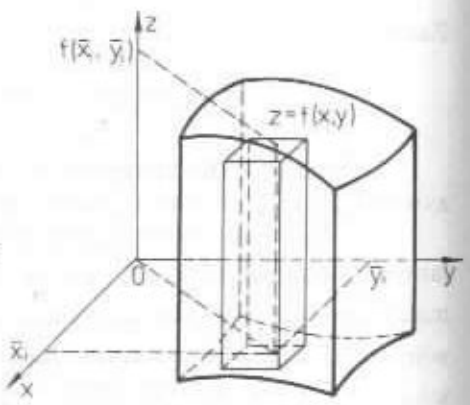
INTERPRETACJA GEOMETRYCZNA. Niech f będzie funkcją c i ą g ł ą i n i e u j e m n ą na regularnym obszarze D. Niech V oznacza bryłę (rys. 1.4) ograniczoną płaszczyzną $z=0$, powierzchnią $z=f(x,y)$ oraz powierzchnią walcową, której kierownicą jest brzeg obszaru D, a tworzące są równoległe do osi 0z, czyli

$$V=\{(x,y,z)\in R^3:\ 0\le z\le f(x,y)\ \wedge\ (x,y)\in D\}.$$

Niżej podamy definicję objętości $|V|$ tej bryły.



Rys 1.4.



Rys 1.5.

Obszar D dzielimy na n obszarów częściowych D_i o polach ΔD_i, $i=1,2,\dots,n$, spełniających warunki 1°, 2° z definicji całki podwójnej. W każdym obszarze D_i obieramy punkt $(\bar{x}_i,\bar{y}_i)$ i rozważamy walce V_i:

$$V_i=\{(x,y,z)\in R^3:\ 0\le z\le f(\bar{x}_i,\bar{y}_i)\ \wedge\ (x,y)\in D_i\},\quad i=1,2,\dots,n,$$

(rys 1.5): Objętość ΔV_i każdego z walców V_i jest równa

$$\Delta V_i=f(\bar{x}_i,\bar{y}_i)\,\Delta D_i,\quad i=1,2,\dots,n.$$

Suma objętości tych walców $\sum_{i=1}^{n}\Delta V_i=\sum_{i=1}^{n}f(\bar{x}_i,\bar{y}_i)\Delta D_i$ jest przybliżoną wartością objętości bryły V. Rozważmy normalny ciąg podziałów obszaru D, odpowiadający mu ciąg $(\sum_{i=1}^{n}\Delta V_i)$, czyli ciąg $(\sum_{i=1}^{n}f(\bar{x}_i,\bar{y}_i)\Delta D_i)$ oraz

granicę tego ciągu

$$\lim_{n\to\infty}\sum_{i=1}^{n}f(\bar{x}_i,\bar{y}_i)\Delta D_i.$$

Wobec ciągłości funkcji f na obszarze D granica ta istnieje i ma tę samą wartość przy każdym normalnym ciągu podziałów i przy każdym wyborze punktów $(\bar{x}_i,\bar{y}_i)$, i jest równa całce podwójnej funkcji f na obszarze D. Naturalnym jest więc przyjąć następującą definicję objętości $|V|$ bryły V:

$$|V|\overset{\text{def}}{=}\lim_{n\to\infty}\sum_{i=1}^{n}\Delta V_i=\lim_{n\to\infty}\sum_{i=1}^{n}f(\bar{x}_i,\bar{y}_i)\Delta D_i,$$

czyli

$$|V|=\iint_D f(x,y)\,dxdy.$$

Zatem całka podwójna funkcji f c i ą g ł e j i n i e u j e m n e j na obszarze regularnym D jest równa objętości bryły V:

$$V=\{(x,y,z)\in R^3:\ 0\le z\le f(x,y)\ \wedge\ (x,y)\in D\}.$$

W szczególności, jeśli funkcja $f(x,y)=1$ dla $(x,y)\in D$, to bryła V jest walcem o wysokości 1 i podstawie D, której objętość jest równa $|V|=|D|\cdot 1=|D|$. Zatem

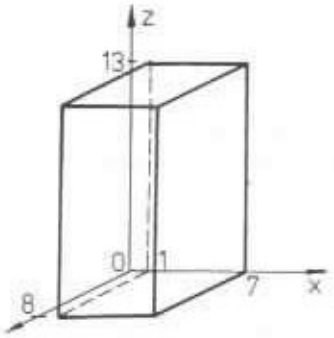
$$\iint_D dxdy=|D|.$$

Dla przykładu: a) całka $\iint_D 13\,dxdy$, jeśli $D=\{(x,y)\in R^2:\ 1\le x\le 7\ \wedge\ 0\le y\le 8\}$, jest równa objętości prostopadłościanu (rys 1.6) o wysokości 13 i podstawie będącej prostokątem o polu równym $6\cdot 8$, zatem

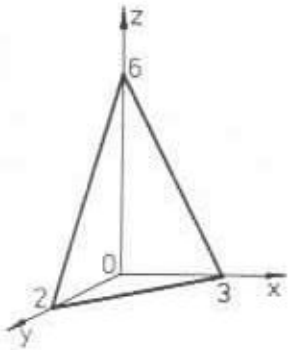
$$\iint_D 13\,dxdy=6\cdot 8\cdot 13=624,$$

b) całka $\iint_D(6-2x-3y)\,dxdy$, jeśli $D=\{(x,y)\in R^2:\ 2x+3y\le 6,\ x,y\ge 0\}$, jest równa objętości $|V|$ ostrosłupa (rys 1.7), zatem

$$\iint_D(6-2x-3y)\,dxdy=|V|=\frac{1}{3}\cdot\frac{1}{2}\cdot 3\cdot 2\cdot 6=6.$$



Rys 1.6.



Rys 1.7.

ZADANIA DO ROZWIĄZANIA.

1. Korzystając z interpretacji geometrycznej obliczyć całki:

a) $\iint_D 4dxdy$, jeśli $D=\{(x,y)\in R^2:\ x^2+y^2\le 9\}$,

b) $\iint_D \sqrt{4x^2+4y^2}dxdy$, jeśli $D=\{(x,y)\in R^2:\ x^2+y^2\le 9\}$,

c) $\iint_D \sqrt{4-y^2}dxdy$, jeśli $D=\{(x,y)\in R^2:\ 0\le x\le 1 \wedge -2\le y\le 2\}$,

d) $\iint_D dxdy$, jeśli $D=\{(x,y)\in R^2:\ x^2+y^2+2x\le 0\}$,

e) $\iint_D dxdy$, jeśli obszar D jest ograniczony krzywymi: $y=x^2$, $y=x^3$.

Odpowiedzi.

1. a) 36π; całka jest równa objętości walca kołowego o wysokości 4 i promieniu 3,
b) 36π; całka jest równa różnicy objętości walca i stożka o promieniu podstawy 3 i wysokości 6,
c) 2π; całka jest równa połowie objętości walca o promieniu 2 i wysokości 1,
d) π; całka jest równa polu koła o promieniu 1,
e) 1/12; całka jest równa polu $|D|=\int_0^1 (x^2-x^3)dx$.

2. WŁASNOŚCI I OBLICZANIE CAŁKI PODWÓJNEJ.

WŁASNOŚCI CAŁKI PODWÓJNEJ. Podamy teraz kilka podstawowych własności całki podwójnej.

(1) *Jeżeli funkcja* f *jest całkowalna na regularnym obszarze* $D\subset R^2$, *to również funkcja* kf, *dla każdej stałej* k, *jest całkowalna oraz*

$$\iint_D kf(x,y)dxdy = k\iint_D f(x,y)dxdy.$$

(2) (addytywność względem funkcji podcałkowej). *Jeżeli funkcje* f *i* g *są całkowalne na regularnym obszarze* $D\subset R^2$, *to również ich suma* f + g *jest funkcją całkowalną na tym obszarze oraz*

$$\iint_D [f(x,y)+g(x,y)]dxdy = \iint_D f(x,y)dxdy + \iint_D g(x,y)dxdy.$$

(3) (addytywność względem obszaru całkowania). *Załóżmy, że regularny obszar* $D\subset R^2$ *jest sumą dwu regularnych obszarów* D_1 *i* D_2 *nie mających wspólnych punktów wewnętrznych. Wówczas funkcja* f *jest całkowalna na obszarze* D *wtedy i tylko wtedy, gdy jest całkowalna na każdym z obszarów* D_1 *i* D_2, *przy czym*

$$\iint_{D_1} f(x,y)dxdy + \iint_{D_2} f(x,y)dxdy = \iint_D f(x,y)dxdy.$$

(4) *Jeżeli funkcja f jest całkowalna na regularnym obszarze D i* $f(x,y)\ge 0$ *dla* $(x,y)\in D$, *to*

$$\iint_D f(x,y)dxdy \ge 0.$$

Krótko: *całka podwójna funkcji nieujemnej jest liczbą nieujemną.*

(5) (monotoniczność całki podwójnej). *Jeżeli funkcje* f *i* g *są całkowalne na regularnym obszarze* $D\subset R^2$ *oraz* $f(x,y)\le g(x,y)$ *dla* $(x,y)\in D$, *to również*

$$\iint_D f(x,y)dxdy \le \iint_D g(x,y)dxdy .$$

(6) (dwustronne oszacowanie całki podwójnej). *Jeżeli* f *jest funkcją całkowalną na regularnym obszarze* $D \subset R^2$ *oraz dla* $(x,y) \in D$, *zachodzi nierówność* $m \le f(x,y) \le M$, *to*

$$m|D| \le \iint_D f(x,y)dxdy \le M|D| .$$

(7) *Zmiana wartości funkcji* f *w punktach leżących na skończonej liczbie wykresów funkcji postaci* $y=y(x)$ *lub* $x=x(y)$ *zawartych w regularnym obszarze* D *nie wpływa na całkowalność tej funkcji w obszarze* D *ani na wartość całki podwójnej, jeśli całka na tym obszarze istnieje.*

Własności te dowodzi się korzystając z definicji całki podwójnej. Dla przykładu udowodnimy własność (5).

D o w ó d (5). Prawdziwość tej własności wynika z następującego ciągu implikacji

$$\begin{array}{rcll}
f(x,y) & \le & g(x,y) & \text{dla } (x,y) \in D \\
\Downarrow & & & \\
f(\bar{x}_i,\bar{y}_i) & \le & g(\bar{x}_i,\bar{y}_i) & \text{dla } (\bar{x}_i,\bar{y}_i) \in D_i,\ i=1,2,\ldots,n \\
\Downarrow & & & \\
f(\bar{x}_i,\bar{y}_i)\Delta D_i & \le & g(\bar{x}_i,\bar{y}_i)\Delta D_i & \text{dla } (\bar{x}_i,\bar{y}_i) \in D_i,\ i=1,2,\ldots,n \\
\Downarrow & & & \\
\sum_{i=1}^{n} f(\bar{x}_i,\bar{y}_i)\Delta D_i & \le & \sum_{i=1}^{n} g(\bar{x}_i,\bar{y}_i)\Delta D_i & \\
\Downarrow & & & \\
\lim_{n\to\infty} \sum_{i=1}^{n} f(\bar{x}_i,\bar{y}_i)\Delta D_i & \le & \lim_{n\to\infty} \sum_{i=1}^{n} g(\bar{x}_i,\bar{y}_i)\Delta D_i & \\
\Downarrow & & & \\
\iint_D f(x,y)dxdy & \le & \iint_D g(x,y)dxdy . & \square
\end{array}$$

ZAMIANA CAŁKI PODWÓJNEJ NA ITEROWANĄ (DWUKROTNĄ). Podamy teraz twierdzenie dotyczące sposobu obliczania całki podwójnej na obszarze normalnym.

TWIERDZENIE 2.1 (o zamianie całki podwójnej na iterowaną). Jeśli funkcja f jest ciągła na obszarze D normalnym względem osi 0x, przy czym

$$D = \{(x,y) \in R^2:\ a \le x \le b \ \wedge\ g(x) \le y \le h(x)\},$$

to

$$\iint_D f(x,y)dxdy = \int_a^b \left[\int_{g(x)}^{h(x)} f(x,y)dy \right] dx .$$

Jeśli zaś obszar D jest normalny względem osi 0y, przy czym

$$D = \{(x,y) \in R^2:\ c \le y \le d \ \wedge\ k(y) \le x \le l(y)\},$$

to

$$\iint_D f(x,y)dxdy = \int_c^d \left[\int_{k(y)}^{l(y)} f(x,y)dx \right] dy .$$

W przypadku szczególnym, gdy obszar D jest prostokątem o bokach równoległych do osi 0x i 0y, przy czym

$$D = \{(x,y) \in R^2:\ a \le x \le b \ \wedge\ c \le y \le d\},$$

to

$$\iint_D f(x,y)dxdy = \int_a^b \left[\int_c^d f(x,y)dy \right] dx = \int_c^d \left[\int_a^b f(x,y)dx \right] dy .$$

WNIOSEK. Jeżeli obszar D jest prostokątem

$$D = \{(x,y) \in R^2:\ a \le x \le b \ \wedge\ c \le y \le d\}$$

oraz funkcja f daje przedstawić się w postaci iloczynu

$$f(x,y) = \varphi(x)\psi(y),$$

przy czym funkcje φ *i* ψ są ciągłe odpowiednio na przedziałach $<a,b>$ i $<c,d>$, to

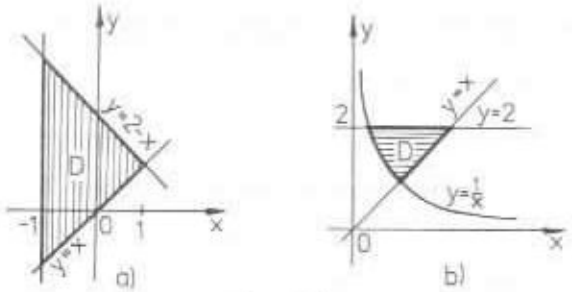
$$\iint_D f(x,y)dxdy = \left(\int_a^b \varphi(x)dx \right) \left(\int_c^d \psi(y)dy \right) .$$

PRZYKŁAD 2.1. Obliczymy całki podwójne:

a) $\iint_D (2y+6x^2)dxdy$, jeśli $D=\{(x,y)\in R^2: x\le y\le 2-x \wedge x\ge -1\}$,

b) $\iint_D 2xy^2dxdy$, jeśli D jest obszarem ograniczonym łukami krzywych: $y=x$, $y=1/x$, $y=2$.

W każdym z tych przykładów obszar D jest obszarem normalnym względem obu osi (rys. 2.1 a) i b)). W punkcie a) obszar D zapiszemy jako normalny względem osi 0x, a w punkcie b) jako normalny względem osi 0y.



Rys 2.1

a) Obszar D zapisujemy nierównościami:

$$D=\{(x,y)\in R^2: -1\le x\le 1 \wedge x\le y\le 2-x\}$$

i obliczamy

$$\iint_D (2y+6x^2)dxdy=\int_{-1}^{1}\left[\int_{x}^{2-x}(2y+6x^2)dy\right]dx=\int_{-1}^{1}\left[y^2+6x^2y\right]_{y=x}^{y=2-x}dx=$$

$$=\int_{-1}^{1}((2-x)^2+6x^2(2-x)-x^2-6x^3)dx=\int_{-1}^{1}(4-4x+12x^2-12x^3)dx=16.$$

b) Ponieważ $D=\{(x,y)\in R^2: 1\le y\le 2 \wedge 1/y\le x\le y\}$, więc

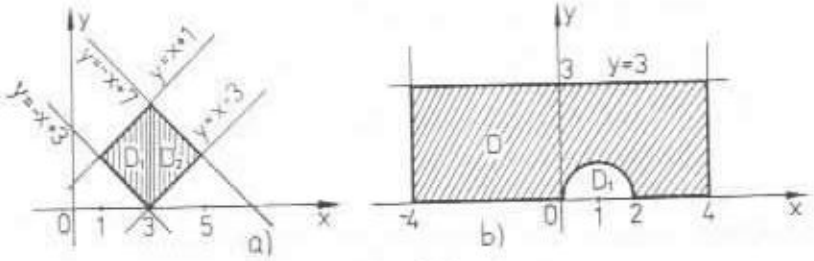
$$\iint_D 2xy^2dxdy=\int_{1}^{2}\left[\int_{1/y}^{y}2xy^2dx\right]dy=\int_{1}^{2}x^2y^2\Big|_{x=1/y}^{x=y}dy=\int_{1}^{2}(y^4-1)dy=\frac{26}{5}.$$ ■

PRZYKŁAD 2.2. Obliczymy następujące całki

a) $\iint_D xdxdy$, jeśli D jest obszarem ograniczonym odcinkami prostych:

$y=x-3$, $y=x+1$, $y=-x+3$, $y=-x+7$,

b) $\iint_D 2xydxdy$, gdy $D=\{(x,y)\in R^2: |x|\le 4 \wedge 0\le y\le 3 \wedge x^2+y^2\ge 2x\}$.



Rys 2.2

a) Obszar D (rys. 2.2 a)) jest normalny względem obu osi, przy czym $D=D_1\cup D_2$, gdzie

$$D_1=\{(x,y)\in R^2: 1\le x\le 3 \wedge -x+3\le y\le x+1\},$$

$$D_2=\{(x,y)\in R^2: 3\le x\le 5 \wedge x-3\le y\le -x+7\}.$$

Ponieważ obszary D_1 i D_2 nie mają wspólnych punktów wewnętrznych, więc

$$\iint_D xdxdy=\iint_{D_1} xdxdy+\iint_{D_2} xdxdy=\int_{1}^{3}\left[\int_{-x+3}^{x+1}xdy\right]dx+\int_{3}^{5}\left[\int_{x-3}^{-x+7}xdy\right]dx=$$

$$=\int_{1}^{3}x(2x-2)dx+\int_{3}^{5}x(-2x+10)dx=\frac{28}{3}+\frac{44}{3}=24.$$

b) Obszar całkowania D przedstawiony jest na rysunku 2.2 b). Przyjmijmy oznaczenia

$$D_1=\{(x,y)\in R^2: 0\le x\le 2 \wedge 0\le y\le \sqrt{2x-x^2}\},$$

$$D_2=\{(x,y)\in R^2: -4\le x\le 4 \wedge 0\le y\le 3\}.$$

Wówczas $D_2=D\cup D_1$. Ponieważ obszary D i D_1 nie mają wspólnych punktów wewnętrznych, więc

$$\iint_{D_2} 2xydxdy = \iint_{D} 2xydxdy + \iint_{D_1} 2xydxdy,$$

a stąd

$$\iint_{D} 2xydxdy = \iint_{D_2} 2xydxdy - \iint_{D_1} 2xydxdy =$$

$$= \int_{-4}^{4}[\int_{0}^{3} 2xydy]dx - \int_{0}^{2}[\int_{0}^{\sqrt{2x-x^2}} 2xydy]dx = \int_{-4}^{4} 9xdx - \int_{0}^{2}(2x^2 - x^3)dx = -\frac{4}{3}.$$

■

PRZYKŁAD 2.3. Obliczymy objętość $|V|$ bryły ograniczonej powierzchniami: $z=0$, $z=x^2+y^2$, $y=x$, $y=2x$, $x=2$.

Rysunek 2.3 przedstawia bryłę V oraz obszar płaski D, który jest prostokątnym rzutem bryły V na płaszczyznę 0xy. Ponieważ

$$V=\{(x,y,z)\in R^3:\ 0\le z\le x^2+y^2 \ \wedge\ (x,y)\in D\},$$

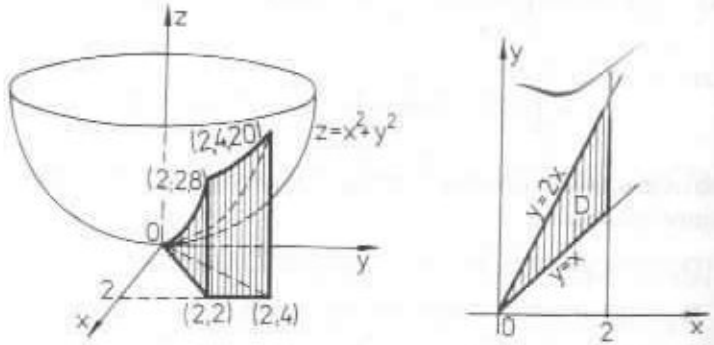
$$D=\{(x,y)\in R^2:\ 0\le x\le 2 \ \wedge\ x\le y\le 2x\},$$

więc

$$|V| = \iint_{D}(x^2+y^2)dxdy = \int_{0}^{2}[\int_{x}^{2x}(x^2+y^2)dy]dx = \int_{0}^{2}[x^2y+\frac{1}{3}y^3]_x^{2x}dx =$$

$$= \int_{0}^{2}(2x^3+\frac{8}{3}x^3-x^3-\frac{1}{3}x^3)dx = \frac{40}{3}.$$

■



Rys 2.3

ZADANIA DO ROZWIĄZANIA.

1. Naszkicować obszar $D\subset R^2$, który jest obszarem normalnym: a) względem obu osi, b) tylko względem osi 0x, c) tylko względem osi 0y.

2. Naszkicować obszar domknięty D ograniczony liniami: $y=x^2+4x$, $y-5=0$, $y+x=0$. Czy jest to obszar normalny względem obu osi układu ?

3. Obliczyć całkę $\iint_D xy^2dxdy$, jeśli D jest obszarem ograniczonym krzywymi:

   a) $x=0$, $x=2$, $y=1$, $y=3$, b) $x=1$, $x=2$, $|y|=1/x$,

   c) $y=1$, $y=2$, $y|x|=1$, d) $y=x^3$, $y=x^2$.

4. Obliczyć całkę $\iint_D xe^ydxdy$, jeśli D jest obszarem ograniczonym krzywymi:

   a) $x=3$, $y=0$, $y=\ln x$, b) $x=1/3$, $y=0$, $y=\ln x$,

   c) $y=0$, $y=x$, $y=2-x$, d) $y=x^2$, $y=x$.

5. Obliczyć całki podwójne:

   a) $\iint_D xdxdy$, jeśli $D=\{(x,y)\in R^2:\ 2x-1\le y\le x \ \wedge\ x\ge 0\}$,

   b) $\iint_D xdxdy$, jeśli $D=\{(x,y)\in R^2:\ -x\le y\le 2x-x^2\}$,

   c) $\iint_D (2y+x)dxdy$, jeśli $D=\{(x,y)\in R^2:\ x^2-1\le y\le 2-2x^2\}$,

   d) $\iint_D e^xdxdy$, jeśli $D=\{(x,y)\in R^2:\ x^2-3x\le y\le x-x^2\}$,

   e) $\iint_D \frac{y}{x}dxdy$, jeśli $D=\{(x,y)\in R^2:\ 0\le y\le \ln x \ \wedge\ x\le e\}$,

f) $\iint_D x\cos y\,dxdy$, jeśli $D=\{(x,y)\in R^2: 0\le y\le 2-x \wedge y\le x\}$,

g) $\iint_D 2xy\,dxdy$, jeśli $D=\{(x,y)\in R^2: x^2+y^2-6x\le 0 \wedge y-x\ge 0\}$,

h) $\iint_D \frac{1}{y+1}dxdy$, jeśli $D=\{(x,y)\in R^2: 0\le y\le 2-x \wedge y-1\le 0 \wedge x+1\ge 0\}$,

i) $\iint_D \frac{1}{x+2}dxdy$, jeśli $D=\{(x,y)\in R^2: x^2\le y\le x+2\}$,

j) $\iint_D \ln y\,dxdy$, jeśli $D=\{(x,y)\in R^2: x+1\le 2y\le 2x \wedge y\le 2\}$,

k) $\iint_D (x+y)dxdy$, jeśli $D=\{(x,y)\in R^2: x^2+y^2\le 1 \wedge x\ge 0\}$,

l) $\iint_D \frac{1}{2+\sqrt{3+y}}dxdy$, jeśli $D=\{(x,y)\in R^2: 2y-2\le x\le 2-2y \wedge y\ge 0\}$,

ł) $\iint_D e^y dxdy$, jeśli $D=\{(x,y)\in R^2: y^2\le x\le y\}$.

6. Obliczyć całki podwójne:

a) $\iint_D \frac{2\ln x}{1+y^2}dxdy$, jeśli D jest obszarem ograniczonym prostymi $x=1$, $x=e$, $y=0$, $y=1$,

b) $\iint_D x\,dxdy$, jeśli D jest obszarem ograniczonym parabolami $y=x^2-6x+8$, $y=4x-x^2$,

c) $\iint_D \frac{4}{y^3}e^{2x}dxdy$, jeśli D jest obszarem ograniczonym krzywymi $y=e^x$, $y=e^{2x}$, $x=1$,

d) $\iint_D \frac{1}{x^4}dxdy$, jeśli D jest obszarem ograniczonym krzywymi $y=x^2-3x+2$, $y=x-1$,

e) $\iint_D \frac{1}{\ln^2 x}dxdy$, jeśli D jest obszarem ograniczonym krzywymi $|y|=\frac{1}{x}$, $x=e$, $x=\sqrt{e}$,

f) $\iint_D xy\,dxdy$, jeśli D jest obszarem ograniczonym krzywymi $y=e^{-x}$, $y=0$, $x=0$, $x=1$,

g) $\iint_D y\ln x\,dxdy$, jeśli D jest czworokątem ograniczonym prostymi $y=x$, $y=2x$, $x=1$, $x=2$,

h) $\iint_D \frac{1}{x+1}dxdy$, jeśli D jest obszarem ograniczonym krzywymi $y=x$, $y=x^2$, $x=2$,

i) $\iint_D \frac{2x}{\sqrt{1+y^3}}dxdy$, jeśli D jest trójkątem o wierzchołkach (0,0), (4,2), (2,2),

j) $\iint_D x(1+y)^2 dxdy$, jeśli D jest trójkątem ograniczonym prostymi $y=0$, $x-y+1=0$, $x+2y-2=0$,

k) $\iint_D \cos(xy)dxdy$, jeśli D jest obszarem ograniczonym krzywymi $2y|x|=\pi$, $y=1$, $y=2$,

l) $\iint_D \frac{x\,dxdy}{y+1}$, jeśli D jest czworokątem o wierzchołkach (−1,0), (2,0), (3,1), 0,1).

7. Obliczyć całki:

a) $\iint_D \frac{dxdy}{x^2+4}$, jeśli D jest obszarem ograniczonym parabolami $y=x^2$, $y=2x^2-4$,

b) $\iint_D xdxdy$, jeśli $D=\{(x,y)\in R^2: x^2-x\le y\le\sqrt{2x}\}$,

c) $\iint_D \frac{e^y}{(4+x)^2}dxdy$, jeśli D jest obszarem ograniczonym krzywymi $y=\ln|x|$, $y=0$, $y=1$,

d) $\iint_D \frac{dxdy}{y^2+4y+5}$, jeśli D jest trójkątem o wierzchołkach (-1,0), (1,-2), (3,0),

e) $\iint_D (x+2y)dxdy$, jeśli D jest obszarem ograniczonym krzywymi $y=|x|$, $y=\frac{|x|}{4}$, $y=\frac{4}{x}$,

f) $\iint_D 2xdxdy$, jeśli $D=\{(x,y)\in R^2: x^2-x\le 2y\le 2-2x \wedge y\le 1\}$,

g) $\iint_D e^y dxdy$, jeśli D jest czworokątem, którego kolejnymi wierzchołkami są punkty (-4,0), (2,0), (0,2), (2,6),

h) $\iint_D ydxdy$, jeśli D jest czworokątem, którego kolejnymi wierzchołkami są punkty (-2,0), (0,1), (1,0), (4,6),

i) $\iint_D xdxdy$, jeśli $D=\{(x,y)\in R^2: 0\le y\le\sqrt{9-x^2}\wedge y\ge x-x^2\}$.

Odpowiedzi.

2. Tak. 3. a)52/3, b)1/3, c) 0, d) 1/88. 4. a) 14/3, b) 10/81, c) 2e−4, d) (3−e)/2.

5. a) 1/6, b) 27/4, c) 16/5, d) 8, e) 1/6, f) 2(1−cos1), g) 27/2, h) ln16−1, i) 15/2−4ln4, j) 1/4, k) 2/3, l) $8\sqrt{3}-40/3$, ł) 3−e.

6. a) π/2, b) 45/2, c) $1+e^{-2}$, d) 4/27, e) 2, f) $(1-3e^{-2})/8$, g) ln16−7/6, h) ln(9/4)−1/2, i) 4, j) 4/5, k) ln4, l) $3(1-\ln\sqrt{2})$.

7. a) 2π−4, b) 28/15, c) ln15−ln(16−e²), d) ln5, e) 32/3, f) 0,

g) $\frac{3}{2}e^2+\frac{1}{2}e^6-8$, h) 35/2; wsk. $\iint_D ydxdy=\iint_{D_2} ydxdy-\iint_{D_1} ydxdy$, gdzie

$D_2=\{(x,y): 0\le y\le 6\wedge y-2\le x\le 1+y/2\}$, $D_1=\{(x,y): 0\le y\le 1\wedge 2y-2\le x\le 1-y\}$.

i) −1/12; wsk. $\iint_D xdxdy=\iint_{D_2} xdxdy-\iint_{D_1} xdxdy$, gdzie

$D_2=\{(x,y): -3\le x\le 3\wedge 0\le y\le\sqrt{9-x^2}\}$, $D_1=\{(x,y): 0\le x\le 1\wedge 0\le y\le x-x^2\}$.

3. ZAMIANA ZMIENNYCH W CAŁCE PODWÓJNEJ.

PRZEKSZTAŁCENIE OBSZARU. Przypomnijmy: **przekształceniem (odwzorowaniem)** zbioru Z **na** zbiór Z' nazywamy przyporządkowanie każdemu elementowi zbioru Z dokładnie jednego elementu zbioru Z', przy czym każdy element zbioru Z' jest obrazem co najmniej jednego elementu zbioru Z. Jeżeli natomiast każdy element zbioru Z' jest obrazem dokładnie jednego elementu zbioru Z mówimy, że to przekształcenie (odwzorowanie) jest **wzajemnie jednoznaczne**.

W dalszym ciągu będziemy rozważać przekształcenia obszaru płaskiego $\Delta=\{(u,v)\}\subset R^2$ w obszar płaski $D=\{(x,y)\}\subset R^2$. Przekształcenia takie można opisać równaniami postaci

$$x=x(u,v),\quad y=y(u,v),\quad (u,v)\in\Delta. \tag{3.1}$$

PRZYKŁAD 3.1. Weźmy pod uwagę prostokąt

$$\Delta=\{(u,v)\in R^2: 0\le u\le 1\wedge -2\le v\le 1\}$$

i przekształcenie

$$x=u+2v,\quad y=u-v. \tag{1}$$

Znajdziemy obraz prostokąta Δ w przekształceniu (1).

Z równań (1) otrzymujemy, że

$$u=\frac{1}{3}(x+2y),\quad v=\frac{1}{3}(x-y).$$

Ponieważ

$$(u,v)\in\Delta \Leftrightarrow (0\le u\le 1 \wedge -2\le v\le 1) \Leftrightarrow$$

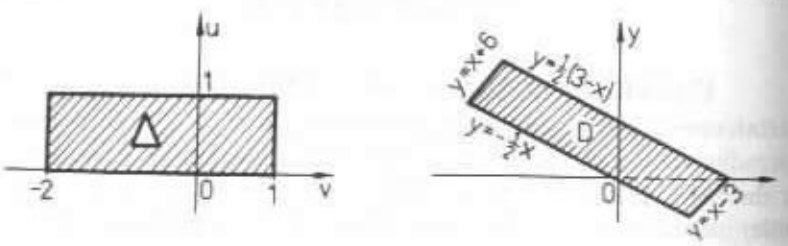
$$\Leftrightarrow \left(0\le\frac{1}{3}(x+2y)\le 1 \wedge -2\le\frac{1}{3}(x-y)\le 1\right) \Leftrightarrow$$

$$\Leftrightarrow \left(-\frac{x}{2}\le y\le\frac{-x+3}{2} \wedge x-3\le y\le x+6\right),$$

więc

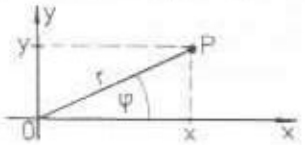
$$(u,v)\in\Delta \Leftrightarrow (x,y)\in D=\left\{(x,y)\in R^2: -\frac{x}{2}\le y\le\frac{-x+3}{2} \wedge x-3\le y\le x+6\right\}.$$

Oznacza to, że obszar domknięty D (równoległobok) jest obrazem prostokąta Δ w przekształceniu (1). Obszary Δ i D są przedstawione na rysunku 3.1. ■



Rys 3.1.

Na płaszczyźnie obierzmy, tak jak na rysunku 3.2, dwa układy



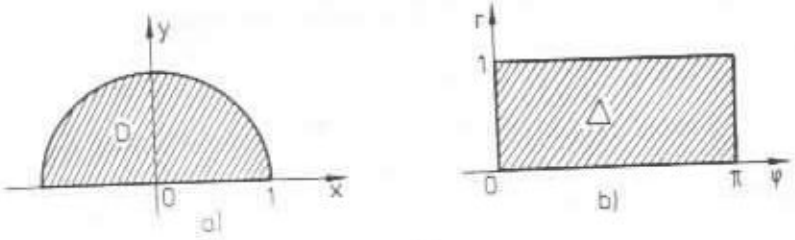
Rys 3.2

współrzędnych: prostokątny 0xy i biegunowy, przy czym oś biegunowa 0r pokrywa się z dodatnią półosią 0x. Dowolnemu punktowi P płaszczyzny przyporządkowujemy współrzędne (x,y) w układzie 0xy oraz współrzędne (r,φ) w układzie biegunowym. Współrzędna r oznacza odległość punktu P od początku układu, zatem $r\ge 0$. Natomiastr φ jest kątem, jaki tworzy wektor o początku w punkcie 0 i końcu w punkcie P z osią biegunową. Ponieważ punkty (r,φ) i $(r,\varphi+2k\pi)$ pokrywają się, więc wystarczy przyjąć, że współrzędna φ przyjmuje wartość z dowolnego

przedziału o długości 2π. Między współrzędnymi (x,y) i (r,φ) zachodzą następujące związki

(3.2) $$x=r\cos\varphi, \quad y=r\sin\varphi.$$

PRZYKŁAD 3.2. Niech $D=\{(x,y)\in R^2: x^2+y^2\le 1 \wedge y\ge 0\}$. Obierzmy układ biegunowy tak, jak na rysunku 3.2 i załóżmy, że współrzędna $\varphi\in(-\pi,\pi>$. Z rysunku 3.3 a) odczytujemy, że punkt $(x,y)\in D$ wtedy i tylko wtedy, gdy współrzędne biegunowe tego punktu spełniają warunki: $0\le r\le 1$, $0\le\varphi\le\pi$.

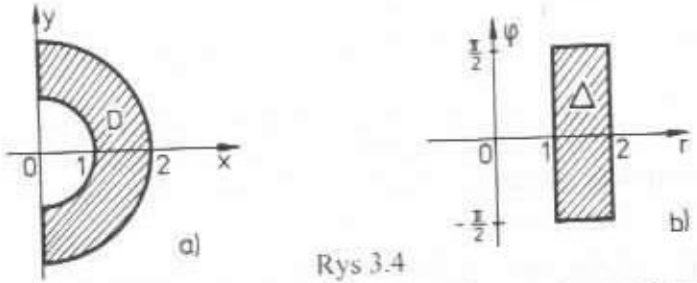


Rys 3.3

Zatem

$$(x,y)\in D \Leftrightarrow (r,\varphi)\in\Delta=\{(r,\varphi): 0\le r\le 1 \wedge 0\le\varphi\le\pi\}.$$

Jeżeli teraz współrzędne biegunowe r i φ potraktujemy jako współrzędne układu prostokątnego $0r\varphi$, to możemy powiedzieć, że obszar D jest obrazem prostokąta Δ (rys. 3.3 b)) w przekształceniu (3.2).



Rys 3.4

Podobnie, korzystając z interpretacji geometrycznej zmiennych r i φ, otrzymujemy, że obszar

$$D=\{(x,y)\in R^2: 1\le x^2+y^2\le 4 \wedge x\ge 0\}$$

(rys 3.4 a)) w przekształceniu (3.2) jest obrazem prostokąta

$$\Delta=\{(r,\varphi)\in R^2:\ 1\le r\le 2\ \wedge\ -\frac{\pi}{2}\le\varphi\le\frac{\pi}{2}\}.$$

Dodajmy, że w obu przykładach odwzorowanie wnętrza obszaru Δ na wnętrze obszaru D jest wzajemnie jednoznaczne. ■

ZAMIANA ZMIENNYCH W CAŁCE PODWÓJNEJ. W pierwszym tomie książki poznaliśmy twierdzenie o zamianie zmiennej w całce oznaczonej (tw. o całkowaniu przez podstawianie). Teraz podamy twierdzenie o zamianie zmiennych w całce podwójnej.

TWIERDZENIE 3.1 (o zamianie zmiennych w całce podwójnej). Jeżeli

1) funkcja $f(x,y)$ jest ciągła na regularnym obszarze D,

2) układ funkcji $x=x(u,v)$, $y=y(u,v)$ przekształca regularny obszar Δ na obszar D, przy czym przekształcenie wnętrza obszaru Δ na wnętrze obszaru D jest wzajemnie jednoznaczne,

3) funkcje $x(u,v)$, $y(u,v)$ są klasy C^1 na pewnym obszarze $\Delta_0\supset\Delta$, przy czym wyznacznik funkcyjny

$$J(u,v)=\begin{vmatrix} x'_u & x'_v \\ y'_u & y'_v \end{vmatrix} \tag{3.3}$$

jest różny od zera wewnątrz obszaru Δ, to

$$\iint_D f(x,y)\,dxdy=\iint_\Delta f(x(u,v),y(u,v))\,|J(u,v)|\,dudv. \tag{3.4}$$

Dodajmy, że wyznacznik funkcyjny (3.3) nazywany **jakobianem** rozważanego przekształcenia.

Uwaga. Wzór (3.4) pozostaje prawdziwy, gdy założenie 2) tego twierdzenia nie jest spełnione w punktach leżących na skończonej liczbie wykresów funkcji postaci $x=x(y)$, $y=y(x)$ zawatych w obszarze D lub $u=u(v)$, $v=v(u)$ zawartych w obszarze Δ.

W dalszym ciągu będziemy korzystać z twierdzenia 3.1 przede wszystkim w przypadku zamiany zmiennych prostokątnych x i y na współrzędne biegunowe r i φ:

$$x=r\cos\varphi,\quad y=r\sin\varphi,$$

gdzie $r\ge 0$, a φ przyjmuje wartości z dowolnego przedziału o długości 2π. Zamiana taka jest szczególnie korzystna, jak się dalej przekonamy, gdy obszar całkowania jest kołem, pierścieniem lub ich wycinkiem, a funkcja podcałkowa jest postaci $f(x,y)=g(x^2+y^2)$. Ponieważ jakobian tego przekształcenia jest równy

$$J(r,\varphi)=\begin{vmatrix} x'_r & x'_\varphi \\ y'_r & y'_\varphi \end{vmatrix}=\begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix}=r,$$

więc wzór (3.4) przyjmuje postać

$$\iint_D f(x,y)\,dxdy=\iint_\Delta f(r\cos\varphi,r\sin\varphi)\,r\,drd\varphi. \tag{3.5}$$

PRZYKŁAD 3.3. Obliczymy całki:

a) $\iint_D\sqrt{x^2+y^2}\,dxdy$, gdy $D=\{(x,y)\in R^2:\ x^2+y^2\le 1\ \wedge\ x\ge 0\}$,

b) $\iint_D(x^2+y^2)\,dxdy$, gdy $D=\{(x,y)\in R^2:\ 4\le x^2+y^2\le 9\ \wedge\ y\ge 0\}$,

c) $\iint_D(x-y^2)\,dxdy$, gdy $D=\{(x,y)\in R^2:\ x^2+4y^2\le 4\ \wedge\ x\ge 0\}$.

a) Obszar D jest obszarem normalnym względem obu osi układu (rys 3.5 a)). Można więc zapisać ten obszar jako normalny względem osi 0y i daną całkę podwójną zamienić na iterowaną:

$$\iint_D\sqrt{x^2+y^2}\,dxdy=\left\{D=\{(x,y):\ -1\le y\le 1\wedge 0\le x\le\sqrt{1-y^2}\}\right\}=$$

$$=\int_{-1}^{1}\Big[\int_0^{\sqrt{1-y^2}}\sqrt{x^2+y^2}\,dx\Big]dy.$$

Obliczenie otrzymanej całki wymaga jednak kłopotliwych rachunków.

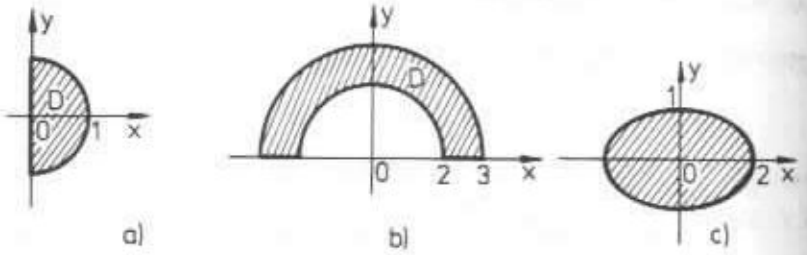
Postąpimy inaczej. Zastosujemy zamianę zmiennych x i y na współrzędne biegunowe r i φ:

$$x=r\cos\varphi,\quad y=r\sin\varphi.$$

Jakobian tego przekształcenia jest równy r. Z rysunku 3.5 a) odczytujemy, że $r\in\langle 0,1\rangle$, $\varphi\in\langle -\frac{\pi}{2},\frac{\pi}{2}\rangle$. Stosując wzór (3.5) otrzymujemy:

$$\iint_D \sqrt{x^2+y^2}\,dxdy = \left\{\begin{array}{l} x=r\cos\varphi,\quad y=r\sin\varphi,\quad J=r \\ \Lambda=\{(r,\varphi):\ 0\le r\le 1\wedge -\frac{\pi}{2}\le\varphi\le\frac{\pi}{2}\}\end{array}\right\} = \iint_\Lambda \sqrt{r^2}\,r\,dr d\varphi =$$

$$= \int_0^1\left[\int_{-\pi/2}^{\pi/2} r^2 d\varphi\right]dr = \int_0^1 \pi r^2 dr = \frac{\pi}{3}.$$



Rys 3.5

b) Obszar D (rys 3.5 b)) jest połową pierścienia o środku (0,0) i promieniach $r_1=2$, $r_2=3$. Całkę obliczymy stosując zamianę zmiennych x i y na współrzędne biegunowe r i φ. Z rysunku łatwo odczytujemy, że $r\in\langle 2,3\rangle$, $\varphi\in\langle 0,\pi\rangle$ i obliczamy:

$$\iint_D (x^2+y^2)dxdy = \left\{\begin{array}{l} x=r\cos\varphi,\quad y=r\sin\varphi,\quad J=r \\ \Lambda=\{(r,\varphi):\ 2\le r\le 3\wedge 0\le\varphi\le\pi\}\end{array}\right\} =$$

$$= \iint_\Delta (r^2\cos^2\varphi + r^2\sin^2\varphi)r\,dr d\varphi = \iint_\Lambda r^3 dr d\varphi = \int_2^3\left[\int_0^\pi r^3 d\varphi\right]dr =$$

$$= \int_2^3 \pi r^3 dr = \frac{65}{4}\pi.$$

c) Obszar całkowania D jest połową elipsy $\frac{x^2}{4}+\frac{y^2}{1}\le 1$ zawartą w półpłaszczyźnie $x\ge 0$ (rys 3.5 c)). W przykładzie tym zastosujemy dwukrotnie zamianę zmiennych:

$$\iint_D (x-y^2)dxdy = \left\{\begin{array}{l} x=2u,\quad y=v,\quad J_1=2 \\ \Lambda_1=\{(u,v):\ u^2+v^2\le 1\wedge u\ge 0\}\end{array}\right\} = \iint_{\Lambda_1}(2u-v^2)2du dv =$$

$$= \left\{\begin{array}{l} u=r\cos\varphi,\quad v=r\sin\varphi,\quad J=r \\ \Lambda=\{(r,\varphi):\ 0\le r\le 1\wedge -\frac{\pi}{2}\le\varphi\le\frac{\pi}{2}\}\end{array}\right\} = 2\iint_\Delta (2r\cos\varphi - r^2\sin^2\varphi)r\,dr d\varphi =$$

$$= 2\int_{-\pi/2}^{\pi/2}\left[\int_0^1 (2r^2\cos\varphi - r^3\sin^2\varphi)dr\right]d\varphi = 2\int_{-\pi/2}^{\pi/2}\left(\frac{2}{3}\cos\varphi - \frac{1}{4}\sin^2\varphi\right)d\varphi =$$

$$= \left[\frac{4}{3}\sin\varphi - \frac{1}{4}\varphi + \frac{1}{8}\sin 2\varphi\right]_{-\pi/2}^{\pi/2} = \frac{8}{3}-\frac{\pi}{4}. \quad \blacksquare$$

PRZYKŁAD 3.4 Obliczymy całki:

a) $\iint_D x dxdy$, gdy $D=\{(x,y)\in R^2:\ x^2+y^2-y\le 0\wedge x\le y\}$.

b) $\iint_D x dxdy$, gdy $D=\{(x,y)\in R^2: y\le x^2+y^2\le 9\wedge(x\ge 0\vee y\ge 0)\}$.

a) Obszar D jest tą częścią koła o środku w punkcie (0,1/2) i promieniu równym 1/2 która jest zawarta w półpłaszczyźnie $y\ge x$ (rys 3.6 a)). Zastosujmy zamianę zmiennych x i y na zmienne r i φ: $x=r\cos\varphi$, $y=r\sin\varphi$. Z rysunku łatwo odczytujemy, że $\varphi\in\langle\pi/4,\pi\rangle$. Zakres zmienności r znajdujemy (uwzględniając, że $\sin\varphi\ge 0$ dla $\varphi\in\langle\pi/4,\pi\rangle$ oraz $r\ge 0$) z równoważności:

$$x^2+y^2-y\le 0 \Leftrightarrow r^2-r\sin\varphi\le 0 \Leftrightarrow r(r-\sin\varphi)\le 0 \Leftrightarrow 0\le r\le\sin\varphi.$$

Mamy więc

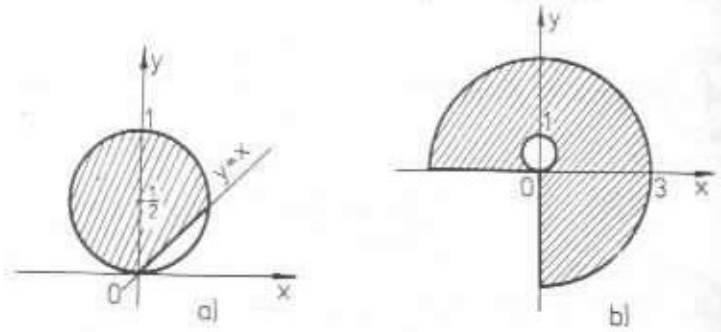
$$\iint_D x dxdy = \left\{\begin{array}{l} x=r\cos\varphi,\quad y=r\sin\varphi,\quad J=r \\ \Lambda=\{(r,\varphi):\ \pi/4\le\varphi\le\pi\wedge 0\le r\le\sin\varphi\}\end{array}\right\} = \iint_\Delta r^2\cos\varphi dr d\varphi =$$

$$= \int_{\pi/4}^{\pi}\left[\int_0^{\sin\varphi} r^2\cos\varphi dr\right]d\varphi = \frac{1}{3}\int_{\pi/4}^{\pi}\sin^3\varphi\cos\varphi\,d\varphi = \frac{1}{12}\sin^4\varphi\Big|_{\pi/4}^{\pi} = -\frac{1}{48}.$$

b) Obszar D przedstawiony jest na rysunku 3.6 b). Przyjmijmy oznaczenia

$$D_1=\{(x,y)\in R^2: x^2+y^2\le y\},$$

$$D_2=\{(x,y)\in R^2: x^2+y^2\le 9 \wedge (x\ge 0 \vee y\ge 0)\}.$$



Rys 3.6

Wówczas $D_2=D\cup D_1$, przy czym obszary D i D_1 nie mają wspólnych punktów wewnętrznych. Zatem

$$\iint_{D_2} x\,dxdy=\iint_{D} x\,dxdy+\iint_{D_1} x\,dxdy$$

a stąd

$$\iint_{D} x\,dxdy=\iint_{D_2} x\,dxdy-\iint_{D_1} x\,dxdy=\left\{\begin{matrix}x=r\cos\varphi,\ y=r\sin\varphi,\\ J=r\end{matrix}\right\}=$$

$$=\iint_{\Delta_2} r^2\cos\varphi\,dr\,d\varphi-\iint_{\Delta_1} r^2\cos\varphi\,dr\,d\varphi=\left\{\begin{matrix}\Delta_2=\{(r,\varphi):-\frac{\pi}{2}\le\varphi\le\pi\wedge 0\le r\le 3\}\\ \Delta_1=\{(r,\varphi):0\le\varphi\le\pi\wedge 0\le r\le\sin\varphi\}\end{matrix}\right\}=$$

$$=\int_{-\pi/2}^{\pi}\left[\int_0^3 r^2\cos\varphi\,dr\right]d\varphi-\int_0^{\pi}\left[\int_0^{\sin\varphi} r^2\cos\varphi\,dr\right]d\varphi=$$

$$=\int_{-\pi/2}^{\pi}9\cos\varphi\,d\varphi-\int_0^{\pi}\frac{1}{3}\sin^3\varphi\cos\varphi\,d\varphi=9\sin\varphi\Big|_{-\pi/2}^{\pi}-\frac{1}{12}\sin^4\varphi\Big|_0^{\pi}=9.\ \blacksquare$$

ZADANIA DO ROZWIĄZANIA

1. Opisać nierównościami oraz naszkicować obszar Δ, który jest obrazem trójkąta D ograniczonego prostymi $y-2x=0$, $2y-x=0$, $x+y=3$ w przekształceniu $u=2y-x$, $v=y-2x$.

2. Opisać nierównościami i naszkicować obszar Δ, który jest obrazem czworokąta $D=\{(x,y)\in R^2: 0\le y\le 2\wedge y-2\le x\le 1\}$ w przekształceniu: $u=y-2$, $v=x-y+1$.

3. Znaleźć obraz Δ obszaru $D=\{(x,y)\in R^2: 4x^2+9y^2\le 36\wedge x\le 0\}$ w przekształceniu $u=2x$, $v=3y$. Obliczyć jakobian tego przekształcenia oraz naszkicować obszary D i Δ.

4. Opisać nierównościami i naszkicować obszar Δ, którego obrazem w przekształceniu $x=r\cos\varphi$, $y=r\sin\varphi$ przy założeniu, że $r\ge 0$ i $0\le\varphi\le 2\pi$, jest obszar D:

   a) $D=\{(x,y)\in R^2: x^2+y^2\le 9\wedge x\ge 0\wedge y\le 0\}$,

   b) $D=\{(x,y)\in R^2: 1\le x^2+y^2\le 9\wedge y\le 0\}$,

   c) $D=\{(x,y)\in R^2: 1\le x^2+y^2\le 4\ \wedge\ -y\le x\le y\}$,

   d) $D=\{(x,y)\in R^2: x^2+y^2\ge 9\}$.

5. Obliczyć całki podwójne:

   a) $\iint_D xy^2\,dxdy$, jeśli $D=\{(x,y)\in R^2: x^2+y^2\le 1\wedge x\ge 0\}$,

   b) $\iint_D x^2\,dxdy$, jeśli $D=\{(x,y)\in R^2: x^2+y^2\le 1\wedge y\ge 0\}$,

   c) $\iint_D x(x^2+y^2)\,dxdy$, jeśli $D=\{(x,y)\in R^2: x^2+y^2\le 1\wedge y\ge x\}$,

   d) $\iint_D y^2\,dxdy$, jeśli $D=\{(x,y)\in R^2: x^2+y^2\le 4\wedge -x\le y\le x\}$,

   e) $\iint_D (x^2-y^2)\,dxdy$, jeśli $D=\{(x,y)\in R^2: x^2+y^2\le 1\wedge -x\le y\le 0\}$,

f) $\iint_D \sqrt{x^2+y^2}\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2\le 4 \wedge (x\ge 0 \vee y\ge 0)\}$,

g) $\iint_D x^2\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2\le 2\}$,

h) $\iint_D (x+y)\,dxdy$, $D=\{(x,y)\in R^2: 1\le x^2+y^2\le 4 \wedge y\ge 0\}$,

i) $\iint_D \frac{2\,dxdy}{\sqrt{5+x^2+y^2}}$, $D=\{(x,y)\in R^2: x^2+y^2\le 4 \wedge y\le x\le -y\}$,

j) $\iint_D \sin\sqrt{x^2+y^2}\,dxdy$, $D=\{(x,y)\in R^2: 4\pi^2\le x^2+y^2\le 9\pi^2 \wedge y\ge 0\}$,

k) $\iint_D \ln(1+x^2+y^2)\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2\le 4 \wedge x\le y\le -x\}$,

l) $\iint_D y^2 e^{x^2+y^2}\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2\le 1 \wedge x\le 0\}$

ł) $\iint_D \ln(1+\sqrt{x^2+y^2})\,dxdy$, $D=\{(x,y)\in R^2: 1\le x^2+y^2\le 9 \wedge x+y\ge 0\}$

6. Obliczyć całki podwójne:

a) $\iint_D x\sqrt{x^2+y^2}\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2-x\le 0\}$,

b) $\iint_D \sqrt{x^2+y^2}\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2-2y\le 0\}$,

c) $\iint_D y\sqrt{x^2+y^2}\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2\le -2x \wedge x\le y\le -x\}$,

d) $\iint_D 2x^3y\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2-2y\le 0 \wedge x+y\ge 0\}$

e) $\iint_D x\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2+y\le 0 \wedge y-x\le 0\}$,

f) $\iint_D (x+y)\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2+x\le 0 \wedge x+y\ge 0\}$,

g) $\iint_D x\,dxdy$, $D=\{(x,y)\in R^2: 4y\le x^2+y^2\le 6y\}$,

h) $\iint_D \sqrt{x^2+y^2}\,dxdy$, $D=\{(x,y)\in R^2: x\le x^2+y^2\le 2x \wedge y\ge 0\}$,

i) $\iint_D xy\,dxdy$, $D=\{(x,y)\in R^2: -y\le x^2+y^2\le -3y \wedge x\ge y\}$,

j) $\iint_D (y-x)\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2-6y+8\le 0 \wedge y\ge 3\}$,

k) $\iint_D (y+1)^2\,dxdy$, $D=\{(x,y)\in R^2: x^2+y^2-4x+2y-11\le 0 \wedge x\ge 2\}$

l) $\iint_D \sqrt{x^2+y^2}\,dxdy$, $D=\{(x,y)\in R^2: 2x\le x^2+y^2\le 9 \wedge (x\ge 0 \vee y\ge 0)\}$,

ł) $\iint_D (4-5xy^2)\,dxdy$, $D=\{(x,y)\in R^2: 4x^2+9y^2+4x\le 36 \wedge 2x\le 3y\}$.

Odpowiedzi.

1. $\Delta=\{(u,v)\in R^2: 0\le u\le 3 \wedge u-3\le v\le 0\}$.
2. $\Delta=\{(u,v)\in R^2: -2\le u\le 0 \wedge -1\le v\le -u\}$.
3. $\Delta=\{(u,v)\in R^2: u^2+v^2\le 36 \wedge u\le 0\}$, $J(x,y)=6$.
4. a) $\Delta=\{(r,\varphi)\in R^2: 0\le r\le 3 \wedge 3\pi/2\le\varphi\le 2\pi\}$,
   b) $\Delta=\{(r,\varphi)\in R^2: 1\le r\le 3 \wedge \pi\le\varphi\le 2\pi\}$,
   c) $\Delta=\{(r,\varphi)\in R^2: 1\le r\le 2 \wedge \pi/4\le\varphi\le 3\pi/4\}$,
   d) $\Delta=\{(r,\varphi)\in R^2: r\ge 3 \wedge 0\le\varphi\le 2\pi\}$.
5. a) 2/15, b) $\pi/8$, c) $-\sqrt{2}/5$, d) $\pi-2$, e) 1/8; f) 4π; g) π, h) 14/3, i) $(3-\sqrt{5})\pi$; j) $5\pi^2$; k) $\frac{\pi}{4}(5\ln 5-4)$; l) $\pi/4$, ł) $\pi(4\ln 4-1)$.
6. a) 4/15, b) 32/9, c) 0, d) 1/10; e) 1/48; f) $(10-3\pi)/96$; g) 0, h) 14/9, i) -5/12, j) $(4+9\pi)/6$, k) 32π, l) $27\pi/2-32/9$, wsk. por. przykład 3.4 b), ł) $12(\pi+\sqrt{2})$.

4. ZASTOSOWANIA GEOMETRYCZNE CAŁKI PODWÓJNEJ.

OBJĘTOŚĆ BRYŁY. Z zagadnieniem tym spotkaliśmy się już w pierwszej części tego rozdziału przy interpretacji geometrycznej całki podwójnej. Uogólnimy teraz uzyskany tam wzór.

TWIERDZENIE 4.1. Jeżeli bryła V jest określona następująco

$$V=\{(x,y,z)\in R^3:\ f_1(x,y)\le z\le f_2(x,y)\wedge (x,y)\in D\}$$

gdzie f_1 *i* f_2 są funkcjami ciągłymi na obszarze regularnym D, to objętość |V| tej bryły wyraża się wzorem

$$|V|=\iint_D [f_2(x,y)-f_1(x,y)]dxdy. \tag{4.1}$$

Wzór (4.1) pozwala obliczyć objętość bryły V ograniczonej powierzchniami $z=f_1(x,y)$, $z=f_2(x,y)$ i powierzchnią walcową, której kierownicą jest brzeg obszaru D, a tworzące są równoległe do osi 0z. Zauważmy, że obszar D jest rzutem prostokątnym bryły V na płaszczyznę 0xy.

PRZYKŁAD 4.1. Obliczymy objętość bryły V ograniczonej powierzchniami:

a) $z=7-x^2$, $z=3$, $y=0$, $y=4$,

b) $z=4-x^2-y^2$, $z=-5$.

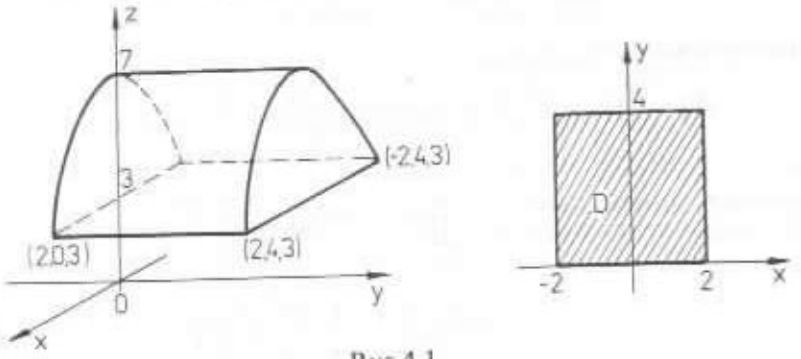
a) Bryłę V i prostokąt D, który jest rzutem prostokątnym tej bryły na płaszczyznę 0xy (rys. 4.1) można zapisać jak następuje:

$$V=\{(x,y,z)\in R^3:\ 3\le z\le 7-x^2\wedge (x,y)\in D\},$$

$$D=\{(x,y)\in R^2:\ -2\le x\le 2\wedge 0\le y\le 4\}.$$

Stąd wynika, że objętość bryły V, zgodnie z (4.1), jest równa

$$|V|=\iint_D (7-x^2-3)dxdy=\int_{-2}^{2}\left[\int_0^4 (4-x^2)dy\right]dx=\int_{-2}^{2}4(4-x^2)dx=\frac{128}{3}.$$



Rys 4.1

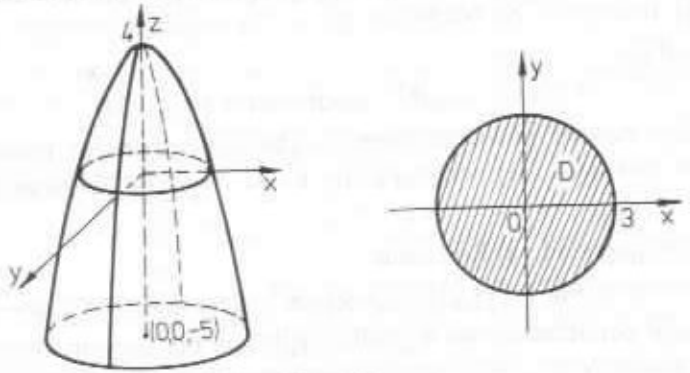
b) Bryła V jest ograniczona paraboloidą $z=4-x^2-y^2$ i płaszczyzną $z=-5$. Krzywą przecięcia tych powierzchni jest okrąg K określony równaniami

$$\begin{cases} z=4-x^2-y^2, \\ z=-5, \end{cases}$$

czyli

$$\begin{cases} 9=x^2+y^2, \\ z=-5. \end{cases}$$

Rzutem prostokątnym bryły V (rys 4.2) na płaszczyznę 0xy jest więc koło D określone nierównością $x^2+y^2\le 9$.



Rys 4.2.

Zatem

$$V=\{(x,y,z)\in R^3:\ -5\le z\le 4-x^2-y^2\wedge (x,y)\in D\},$$

$$D=\{(x,y)\in R^2:\ x^2+y^2\le 9\},$$

a stąd, zgodnie z (4.1), mamy

$$|V|=\iint_D(4-x^2-y^2-(-5))dxdy=\iint_D(9-x^2-y^2)dxdy=$$

$$=\left\{\begin{matrix}x=r\cos\varphi,\ y=r\sin\varphi,\ J=r\\ \Delta=\{(r,\varphi):\ 0\le r\le 1\wedge 0\le\varphi\le 2\pi\}\end{matrix}\right\}=\iint_\Delta(9-r^2)r\,drd\varphi=\int_0^3\int_0^{2\pi}(9r-r^3)d\varphi=$$

$$=2\pi\left[\frac{9}{2}r^2-\frac{1}{4}r^4\right]_0^3=\frac{81}{2}\pi.\ \blacksquare$$

POLE PŁATA POWIERZCHNIOWEGO. Wykres funkcji $z=f(x,y)$ ciągłej na regularnym obszarze $D\subset R^2$ nazywamy **płatem powierzchniowym**. Płat powierzchniowy jest więc zbiorem postaci

$$S=\{(x,y,z)\in R^3:\ z=f(x,y)\wedge(x,y)\in D\}.$$

Jeżeli ponadto pochodne f'_x, f'_y są ciągłe na obszarze D, to płat S nazywamy **płatem powierzchniowym gładkim względem płaszczyzny 0xy**.

Analogicznie: zbiór

$$S=\{(x,y,z)\in R^3:\ y=g(x,z)\wedge(x,z)\in D\}$$

nazywamy płatem powierzchniowym gładkim względem płaszczyzny 0xz, gdy funkcja g i jej pochodne g'_x, g'_z są ciągłe na obszarze regularnym D; zbiór

$$S=\{(x,y,z)\in R^3:\ x=h(y,z)\wedge(y,z)\in D\}$$

nazywamy płatem powierzchniowym gładkim względem płaszczyzny 0yz, gdy funkcja h i jej pochodne h'_y, h'_z są ciągłe na obszarze regularnym D.

Na przykład: a) Równanie

$z=2x+y-6$ dla $(x,y)\in D=\{(x,y)\in R^2:\ 0\le x\le 3\wedge 0\le y\le 6-2x\}$

określa płat powierzchniowy S gładki względem płaszczyzny 0xy. Jest to ta część płaszczyzny $z=2x+y-6$, której rzutem prostokątnym na płaszczyznę 0xy jest trójkąt określony nierównościami: $0\le x\le 3$, $0\le y\le 6-2x$, czyli

$$S=\{(x,y,z)\in R^3:\ z=2x+y-6\wedge 0\le x\le 3\wedge 0\le y\le 6-2x\}.$$

b) Wykres funkcji $x=1+y^2+z^2$, $(y,z)\in D$, gdzie

$$D=\{(y,z)\in R^2:\ y^2+z^2\le 1\}$$

jest płatem powierzchniowym gładkim względem płaszczyzny 0yz. Jest to część paraboloidy $x=1+y^2+z^2$ wycięta walcem o równaniu $y^2+z^2=1$. Płat ten można opisać następująco:

$$S=\{(x,y,z)\in R^3:\ x=1+y^2+z^2\wedge y^2+z^2\le 1\}.$$

c) Wykres funkcji $x=1+\sqrt{y^2+z^2}$, $(y,z)\in D$, gdzie

$$D=\{(y,z)\in R^2:\ y^2+z^2\le 1\}$$

jest płatem powierzchniowym, który jest częścią stożka $x=1+\sqrt{y^2+z^2}$ wyciętą walcem $y^2+z^2=1$. Nie jest to jednak płat gładki, gdyż nie istnieją pochodne x'_y, x'_z w punkcie (0,0).

TWIERDZENIE 4.2. Pole $|S|$ płata powierzchniowego S gładkiego względem płaszczyzny 0xy określonego równaniem

$$z=f(x,y),\ (x,y)\in D,$$

wyraża się wzorem

$$|S|=\iint_D\sqrt{1+(f'_x(x,y))^2+(f'_y(x,y))^2}\,dxdy. \tag{4.2}$$

Analogicznie: pole płata powierzchniowego S gładkiego względem płaszczyzny 0xz określonego równaniem

$$y=g(x,z),\ (x,z)\in D,$$

wyraża się wzorem

$$|S|=\iint_D\sqrt{1+(g'_x(x,z))^2+(g'_z(x,z))^2}\,dxdz. \tag{4.3}$$

PRZYKŁAD 4.2. Obliczymy pole płata powierzchniowego S, gdy

a) S jest częścią stożka $z=\sqrt{x^2+y^2}$, której rzutem na płaszczyznę 0xy jest obszar $D=\{(x,y)\in R^2:\ 4\le x^2+y^2\le 9\wedge -x\le y\le x\}$.

b) S jest częścią paraboloidy $z=1+x^2+y^2$ wyciętą walcem o równaniu $x^2+y^2=2$.

a) Jest to płat gładki względem płaszczyzny 0xy, więc jego pole obliczymy korzystając ze wzoru (4.2). Ponieważ

$$z'_x=\frac{x}{\sqrt{x^2+y^2}},\quad z'_y=\frac{y}{\sqrt{x^2+y^2}},\quad 1+(z'_x)^2+(z'_y)^2=2,$$

więc

$$|S|=\iint_D\sqrt{1+(z'_x)^2+(z'_y)^2}\,dxdy=\iint_D\sqrt{2}\,dxdy=\sqrt{2}\,|D|.$$

Obszar D jest wycinkiem o kącie środkowym $\pi/2$ pierścienia o promieniach 2 i 3. Pole obszaru D jest równe 1/4 pola pierścienia, czyli

$$|D|=\frac{1}{4}(9\pi-4\pi)=\frac{5}{4}\pi.$$

W konsekwencji pole rozważanego płata powierzchniowego jest równe $|S|=5\sqrt{2}\pi/4$.

b) Płat S jest określony równaniem

$$z=1+x^2+y^2,\ \text{gdzie}\ x^2+y^2\le 2,$$

i może być zapisany następująco

$$S=\{(x,y,z)\in R^3:\ z=1+x^2+y^2\ \wedge\ (x,y)\in D\},$$

przy czym

$$D=\{(x,y)\in R^2:\ x^2+y^2\le 2\}.$$

Jest to płat powierzchniowy gładki względem płaszczyzny 0xy.

Pole tego płata obliczymy korzystając ze wzoru (4.2). Ponieważ

$$z'_x=2x,\quad z'_y=2y,\quad 1+(z'_x)^2+(z'_y)^2=1+4x^2+4y^2,$$

więc

$$|S|=\iint_D\sqrt{1+(z'_x)^2+(z'_y)^2}\,dxdy=\iint_D\sqrt{1+4x^2+4y^2}\,dxdy=$$

$$=\left\{\begin{matrix}x=r\cos\varphi,\quad y=r\sin\varphi,\quad J=r\\ \Delta=\{(r,\varphi):0\le r\le\sqrt{2}\ \wedge\ 0\le\varphi\le 2\pi\}\end{matrix}\right\}=\iint_\Delta\sqrt{1+4r^2}\,rdrd\varphi=\int_0^{\sqrt{2}}\Big[\int_0^{2\pi}r\sqrt{1+4r^2}\,d\varphi\Big]dr=$$

$$=2\pi\int_0^{\sqrt{2}}r\sqrt{1+4r^2}\,dr=\frac{\pi}{6}\Big[(\sqrt{1+4r^2})^3\Big]_0^{\sqrt{2}}=\frac{13}{3}\pi.\ \blacksquare$$

PRZYKŁAD 4.3. Obliczymy pole płata powierzchniowego, który jest częścią walca o równaniu $y=x^2$ wyciętą płaszczyznami $z=x$, $z=2x$, $x=2$.

Płat S jest płatem gładkim względem płaszczyzny 0xz, a jego rzutem prostokątnym na płaszczyznę 0xz jest trójkąt D:

$$D=\{(x,z)\in R^2:\ 0\le x\le 2\ \wedge\ x\le z\le 2x\}.$$

Zatem

$$S=\{(x,yz)\in R^3:\ y=x^2\ \wedge\ (x,z)\in D\}.$$

Pole |S| płata S obliczymy korzystając ze wzoru (4.3). Ponieważ

$$y'_x=2x,\quad y'_z=0,\quad 1+(y'_x)^2+(y'_z)^2=1+4x^2,$$

a więc

$$|S|=\iint_D\sqrt{1+(y'_x)^2+(y'_z)^2}\,dxdz=\iint_D\sqrt{1+4x^2}\,dxdz=\int_0^2\Big[\int_x^{2x}\sqrt{1+4x^2}\,dz\Big]dx=$$

$$=\int_0^2 x\sqrt{1+4x^2}\,dx=\frac{1}{12}\Big[(\sqrt{1+4x^2})^3\Big]_0^2=\frac{1}{12}(17\sqrt{17}-1).\ \blacksquare$$

ZADANIA DO ROZWIĄZANIA.

1 Obliczyć objętość bryły V, jeśli

a) $V=\{(x,y,z)\in R^3:\ 0\le z\le x^2+y^2\ \wedge\ y\le 2\ \wedge\ x\le y\le 2x\}$,

b) $V=\{(x,y,z)\in R^3:\ 0\le z\le 1+x^2+y^2\ \wedge\ |x|\le 2-y\ \wedge\ |y|\le 1\}$,

c) $V=\{(x,y,z)\in R^3:\ 0\le z\le 4-x^2\ \wedge\ 0\le y\le 3\}$,

d) $V=\{(x,y,z)\in R^3:\ 0\le 2z\le 3-y\ \wedge\ y\ge x^2-1\}$,

e) $V=\{(x,y,z)\in R^3:\ x^2+y^2-3\le z\le 1+\sqrt{x^2+y^2}\ \wedge\ x^2+y^2\le 1\}$,

f) $V=\{(x,y,z)\in R^3:\ -1\le z\le 1+\sqrt{x^2+y^2}\ \wedge\ x^2+y^2\le y\}$,

g) $V=\{(x,y,z)\in R^3:\ 4x^2+4y^2\le z\le 16\ \wedge\ x^2+y^2+x\le 0\}$,

h) $V=\{(x,y,z)\in R^3:\ 0\le z\le x^2+y^2+1\ \wedge\ x^2+y^2\le 4\ \wedge\ x\le y\le\sqrt{3}x\}$

i) $V=\{(x,y,z)\in R^3:\ x^2+y^2\le z\le 4\ \wedge\ y\ge|x|\}$,

j) $V=\{(x,y,z)\in R^3:\ 9\le z\le x^2+y^2\le 16\}$,

k) $V=\{(x,y,z)\in R^3:\ 0\le z\le 3+\sin y\ \wedge\ |x|\le 1\ \wedge\ |y|\le\pi\}$,

l) $V=\{(x,y,z)\in R^3:\ |z|\le e^x\ \wedge\ |y|\le 1-x\le 1\}$.

2. Obliczyć objętość bryły V ograniczonej powierzchniami:

a) $z=3+x^2$, $z=0$, $y=x-x^2$, $y=-x$,

b) $x+z-4=0$, $z=0$, $y=\sqrt{x}$, $y=2\sqrt{x}$,

c) $x+y+z-2=0$, $z=0$, $y=0$, $y=\sqrt{x}$,

d) $z=x^2+y^2-4$, $z=0$,

e) $z=y^2$, $z=-1$, $y=0$, $y=2-x$, $y=2+x$,

f) $z=2+x^2+y^2$, $z=0$, $x^2+y^2=4$,

g) $z=-3\sqrt{x^2+y^2}$, $z=0$, $x^2+y^2+y=0$,

h) $z=5x^2+5y^2-4$, $z=\sqrt{x^2+y^2}$,

i) $z=y$, $z=-y$, $x^2+y^2-4y=0$,

j) $z=10-x^2-9y^2$, $z=1$,

k) $z=x^2+y^2$, $z=4$, $x^2+y^2=9$, $x^2+y^2=16$,

l) $z=x^2+y^2$, $z=20$, $x^2+y^2=9$, $x^2+y^2=16$.

3. Obliczyć pole płata powierzchniowego S, gdy

a) płat S jest częścią płaszczyzny $z=2x+2y-13$ wyciętą walcem $x^2+y^2-2x=0$,

b) $S=\{(x,y,z)\in R^3:\ z=2x+2y-13\ \wedge\ x^2+y^2\le y\ \wedge\ y\ge|x|\}$,

c) $S=\{(x,y,z)\in R^3:\ z=\sqrt{25-x^2-y^2}\ \wedge\ x^2+y^2\le 16\}$,

d) płat S jest tą częścią paraboloidy $z=9-x^2-y^2$, której rzutem prostokątnym na płaszczyznę 0xy jest pierścień określony nierównością $1\le x^2+y^2\le 2$,

e) płat S jest częścią stożka $z=1+\sqrt{x^2+y^2}$ wyciętą walcem określonym równaniem $x^2+y^2-4y+3=0$,

f) płat S jest wykresem funkcji $z=5+x^2$, $(x,y)\in D$, gdzie

$$D=\{(x,y)\in R^2:\ |y|\le x\ \wedge\ x^2<2\},$$

g) płat S jest tą częścią stożka $z=1+\sqrt{x^2+y^2}$, której rzutem prostokątnym na płaszczyznę 0xy jest obszar

$$D=\{(x,y)\in R^2:\ 1\le x^2+y^2\le 9\ \wedge\ y\le|x|\},$$

h) płat S jest określony równaniem $z=\sqrt{2xy}$, $(x,y)\in D$, gdzie

$$D=\{(x,y)\in R^2:\ 1\le x\le 4\ \wedge\ 4\le y\le 9\},$$

i) płat S jest częścią powierzchni $z=\sqrt{2xy}$ wyciętą płaszczyznami $y=1$, $y=x$, $x=4$,

j) $S=\{(x,y,z)\in R^3:\ y=2xz\ \wedge\ 2\le x^2+z^2\le 6\}$,

k) płat S jest tą częścią stożka $y=1+\sqrt{x^2+z^2}$, której rzutem prostokątnym na płaszczyznę 0xz jest czworokąt

$$D=\{(x,z)\in R^2:\ 1\le x\le 2\ \wedge\ 2x\le z\le 3x\},$$

l) płat S jest określony równaniem $y=\sqrt{z^2-x^2}$, $(x,z)\in D$, gdzie

$$D=\{(x,z)\in R^2:\ 1\le x\le 2\ \wedge\ 2x\le z\le 3x\}.$$

Odpowiedzi.

1. a) $\frac{19}{6}$, b) 24, c) 32, d) $\frac{128}{15}$, e) $\frac{25\pi}{6}$, f) $\frac{9\pi+8}{18}$, g) $\frac{29\pi}{8}$, h) $\frac{\pi}{2}$, i) 2π, j) $\frac{49\pi}{2}$, k) 12π, l) $4(e-2)$.

2. a) $\frac{28}{5}$, b) $\frac{128}{15}$, c) $\frac{17}{20}$, d) 8π, e) $\frac{20}{3}$, f) 16π, g) $\frac{4}{3}$, h) $\frac{13\pi}{6}$, i) 16π, j) $\frac{27\pi}{2}$, k) $\frac{119\pi}{2}$, l) $\frac{105\pi}{2}$.

3. a) 3π, b) $\frac{3}{8}(\pi+2)$, c) 20π, d) $\frac{\pi}{6}(27-5\sqrt{5})$, e) $\sqrt{2}\pi$, f) $\frac{13}{3}$, g) $6\sqrt{2}\pi$, h) $\frac{52\sqrt{2}}{3}$, i) $\frac{28}{3\sqrt{2}}$, j) $\frac{49\pi}{3}$, k) $\frac{3\sqrt{2}}{2}$, l) $6-\frac{3}{2}\sqrt{6}$.

5. PRZYKŁADY ZASTOSOWANIA CAŁKI PODWÓJNEJ W FIZYCE.

W przykładach pokażemy wykorzystanie całki podwójnej do obliczania: masy, momentów statycznych, współrzędnych środka ciężkości i momentów bezwładności obszaru płaskiego.

PRZYKŁAD 5.1. Obliczymy masę prostokątnego arkusza blachy, której gęstość powierzchniowa ρ w punkcie (x,y) jest równa x^2y, gdy brzegiem tego arkusza blachy są odcinki prostych: $x=-1$, $x=2$, $y=0$, $y=4$.

Masa m obszaru D jest równa

$$m = \iint_D \rho(x,y)\,dxdy, \tag{5.1}$$

gdzie $\rho(x,y)$ jest gęstością masy w punkcie $(x,y)\in D$.

Mamy więc

$$m = \iint_D x^2y\,dxdy = \int_{-1}^{2}\Big[\int_0^4 x^2y\,dy\Big]dx = \int_{-1}^{2} x^2dx\cdot\int_0^4 y\,dy = 24. \quad \blacksquare$$

Warto zauważyć, że wzór (5.1) dla obszaru płaskiego jednorodnego D, to znaczy o gęstości stałej ρ w każdym punkcie tego obszaru, sprowadza się do znanego wzoru na masę: $m=\rho\cdot|D|$.

PRZYKŁAD 5.2. Obliczymy masę obszaru

$$D=\{(x,y)\in R^2: y^2\le x\le 1\},$$

gdy gęstość powierzchniowa określona jest wzorem $\rho(x,y)=x^2+y^2$.

Ze wzoru (5.1) mamy

$$m = \iint_D (x^2+y^2)\,dxdy = \int_{-1}^{1}\Big[\int_{y^2}^{1}(x^2+y^2)\,dx\Big]dy = \frac{88}{105}. \quad \blacksquare$$

PRZYKŁAD 5.3. Obliczymy momenty statyczne M_x i M_y odpowiednio względem osi 0x i 0y obszaru D z przykładu 5.2.

Moment statyczny M_x względem osi 0x obszaru D jest równy

$$M_x = \iint_D y\rho(x,y)\,dxdy. \tag{5.2}$$

Mamy więc po uwzględnieniu danych z przykładu 5.2

$$M_x = \iint_D y(x^2+y^2)\,dxdy = \int_{-1}^{1}\Big[\int_{y^2}^{1}(yx^2+y^3)\,dx\Big]dy = 0.$$

Analogicznie obliczamy moment statyczny M_y względem osi 0y

$$M_y = \iint_D x\rho(x,y)\,dxdy = \int_{-1}^{1}\Big[\int_{y^2}^{1}(x^3+xy^2)\,dx\Big]dy = \frac{40}{63}. \quad \blacksquare$$

PRZYKŁAD 5.4. Wyznaczymy współrzędne środka ciężkości obszaru D z przykładu 5.2.

Współrzędne x_s, y_s środka ciężkości obszaru D są równe

$$x_s=\frac{M_y}{m}, \quad y_s=\frac{M_x}{m}. \tag{5.3}$$

Z przykładu 5.2 i 5.3 mamy: $m=88/105$, $M_x=0$, $M_y=40/63$. Zatem środkiem ciężkości obszaru D jest punkt $(25/33,0)$. $\blacksquare$

PRZYKŁAD 5.5. Obliczymy momenty bezwładności B_x i B_y odpowiednio względem osi 0x i 0y obszaru $D=\{(x,y): x^2+y^2\le 9\}$, gdy gęstość powierzchniowa $\rho(x,y)=|x|$.

Moment bezwładności B_x względem osi 0x obszaru D jest równy

$$B_x = \iint_D y^2\rho(x,y)\,dxdy. \tag{5.4}$$

Zatem

$$B_x = \iint_D y^2|x|dxdy = \left\{\begin{matrix} x=r\cos\varphi, & y=r\sin\varphi, & J=r \\ \Delta=\{(r,\varphi):\ 0\le r\le 3 \wedge 0\le\varphi\le 2\pi\} \end{matrix}\right\} =$$

$$= \iint_D r^4\sin^2\varphi\,|\cos\varphi|\,dr\,d\varphi = \int_0^{2\pi}\left[\int_0^3 r^4\sin^2\varphi\,|\cos\varphi|\,dr\right]d\varphi =$$

$$= \frac{243}{5}\int_0^{2\pi}\sin^2\varphi|\cos\varphi|d\varphi =$$

$$= \frac{243}{5}\left[\int_0^{\pi/2}\sin^2\varphi\cos\varphi\,d\varphi + \int_{\pi/2}^{3\pi/2}-\sin^2\varphi\cos\varphi\,d\varphi + \int_{3\pi/2}^{2\pi}\sin^2\varphi\cos\varphi d\varphi\right] =$$

$$= \frac{243}{5}\left(\frac{1}{3}+\frac{2}{3}+\frac{1}{3}\right) = 64{,}8\,.$$

Analogicznie obliczamy moment bezwładności B_y względem osi 0y

$$B_y = \iint_D x^2\rho(x,y)dxdy = \iint_D x^2|x|dxdy = 129{,}6\,. \quad \blacksquare$$

ZADANIA DO ROZWIĄZANIA.

1. Obliczyć masę obszaru $D=\{(x,y)\in R^2:\ x^2+y^2\le 4y\}$, którego gęstość powierzchniowa jest równa $\rho(x,y)$, jeśli:
   a) $\rho(x,y)=y$, b) $\rho(x,y)=x^2$, c) $\rho(x,y)=x^2+y^2$, d) $\rho(x,y)=|x|$.
2. Obliczyć masę obszaru D ograniczonego łukami parabol $y=x^2$, $y^2=x$, którego gęstość powierzchniowa jest równa $\rho(x,y)$, jeśli:
   a) $\rho(x,y)=x$, b) $\rho(x,y)=y$, c) $\rho(x,y)=|x-y|$, d) $\rho(x,y)=\sqrt{x}$.
3. Obliczyć momenty statyczne M_x i M_y obszaru D z zadania 1 a) i b).
4. Obliczyć współrzędne x_s, y_s środka ciężkości obszaru D z zadania 1 a) i b).
5. Obliczyć momenty bezwładności B_x, B_y obszaru D z zadania 1 a).

Odpowiedzi.

1. a) $m=8\pi$, b) $m=4\pi$, c) $m=24\pi$, d) $m=32/3$.
2. a) $m=3/20$, b) $m=3/20$,
   c) $m=1/30$; $m=\int_0^1\left[\int_{x^2}^{\sqrt{x}}|x-y|dy\right]dx=\int_0^1\left[\int_{x^2}^{x}(x-y)dy+\int_{x}^{\sqrt{x}}(y-x)dy\right]dx$,
   d) $m=3/14$.
3. a) $M_x=20\pi$, $M_y=0$, b) $M_x=8\pi$, $M_y=0$.
4. a) $(0,5/2)$, b) $(0,2)$.
5. a) $B_x=56\pi$, $B_y=8\pi$.

6. CAŁKA KRZYWOLINIOWA NIESKIEROWANA.

ŁUK GŁADKI. Zbiór

$$K=\{(x,y)\in R^2:\ x=x(t)\wedge y=y(t)\wedge t\in\langle\alpha,\beta\rangle\}$$

nazywamy **krzywą płaską**. Równania

$$x=x(t),\quad y=y(t),\quad t\in\langle\alpha,\beta\rangle \tag{6.1}$$

nazywamy **równaniami parametrycznymi** tej krzywej.

Ta sama krzywa może mieć różne równania parametryczne. Na przykład: $x=t$, $y=t$, $t\in\langle 0,1\rangle$ oraz $x=t^3$, $y=t^3$, $t\in\langle 0,1\rangle$ są równaniami tego samego odcinka prostej $y=x$ dla $x\in\langle 0,1\rangle$.

Łukiem gładkim (regularnym) nazywamy krzywą K daną równaniami parametrycznymi (6.1), jeżeli

- różnym wartościom parametru $t\in(\alpha,\beta)$ odpowiadają różne punkty (x,y) tej krzywej,
- funkcje x(t), y(t) są klasy C^1 na przedziale $\langle\alpha,\beta\rangle$,
- pochodne tych funkcji nie zerują się jednocześnie, tzn. $(x'(t))^2+(y'(t))^2>0$ dla $t\in\langle\alpha,\beta\rangle$.

Geometrycznie oznacza to, że łuk gładki nie ma punktów wielokrotnych i ma w każdym punkcie styczną zmieniającą się w sposób ciągły przy zmianie punktu styczności.

Krzywą daną równaniami parametrycznymi (6.1), w których funkcje x(t), y(t) są ciągłe na przedziale $<\alpha,\beta>$, nazywamy **krzywą kawałkami gładką**, jeśli nie ma ona punktów wielokrotnych i daje się podzielić na skończoną liczbę łuków gładkich.

Punkty $(x(\alpha),y(\alpha))$, $(x(\beta),y(\beta))$ krzywej danej równaniami parametrycznymi (6.1) nazywamy **końcami** tej krzywej. Jeśli te punkty pokrywają się, to krzywą nazywamy **zamkniętą**.

Krzywa K, która jest wykresem funkcji klasy C^1

$$y=y(x), \quad x\in<a,b>,$$

jest łukiem gładkim, gdyż równania parametryczne tej krzywej możemy zapisać w postaci

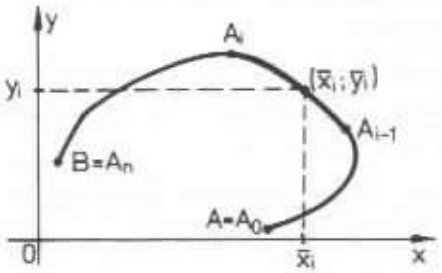
$$x=t, \quad y=y(t), \quad t\in<a,b>,$$

a wtedy łatwo widać, że różnym wartościom parametru t odpowiadają różne punkty krzywej K, funkcje x(t), y(t) są klasy C^1 oraz

$$(x'(t))^2+(y'(t))^2=1+(y'(t))^2>0.$$

Dla przykładu: a) krzywa K: $x=t$, $y=\sqrt{t}$, $t\in<1,4>$ jest łukiem gładkim (co to za krzywa?), b) krzywa K: $x=t$, $y=|t|$, $t\in<-1,4>$ nie jest łukiem gładkim (funkcja $y=|t|$ nie ma pochodnej w punkcie $t=0$; co to za krzywa?); krzywa ta jest sumą dwóch łuków gładkich K_1: $x=t$, $y=t$, $t\in<0,4>$ oraz K_2: $x=t$, $y=-t$, $t\in<-1,0>$; krzywa K jest krzywą kawałkami gładką.

OKREŚLENIE CAŁKI KRZYWOLINIOWEJ NIESKIEROWANEJ. Niech f będzie funkcją określoną na łuku gładkim K o końcach A i B. Podzielmy łuk K punktami $A=A_0,A_1,A_2,\ldots,A_{n-1},A_n=B$ na n łuków częściowych $l_1, l_2,\ldots,l_n$, gdzie $l_i=A_{i-1}A_i$, $i=1,2,\ldots,n$, (rys 6.1).



Rys 6.1.

Niech Δl_i oznacza długość łuku częściowego l_i, $i=1,2,\ldots,n$. Liczbę $\delta_n=\max(\Delta l_1,\Delta l_2,\ldots,\Delta l_n)$ nazywamy średnicą danego podziału.

Na każdym łuku częściowym l_i obierzmy punkt pośredni $(\bar{x}_i,\bar{y}_i)$ i utwórzmy sumę

$$S_n=\sum_{i=1}^{n} f(\bar{x}_i,\bar{y}_i)\Delta l_i.$$

Utwórzmy następnie normalny ciąg podziałów łuku K na łuki częściowe, to znaczy taki ciąg, że $\delta_n\to 0$ przy $n\to\infty$. Ciągowi temu odpowiada ciąg (S_n). Rozważmy granicę

$$\lim_{n\to\infty} S_n=\lim_{n\to\infty}\sum_{i=1}^{n} f(\bar{x}_i,\bar{y}_i)\Delta l_i.$$

Jeżeli dla każdego normalnego ciągu podziałów łuku K i każdego wyboru punktów pośrednich $(\bar{x}_i,\bar{y}_i)$ istnieje ta sama skończona granica ciągu (S_n), to tę granicę nazywamy **całką krzywoliniową nieskierowaną (niezorientowaną)** funkcji f po łuku K i oznaczamy symbolem

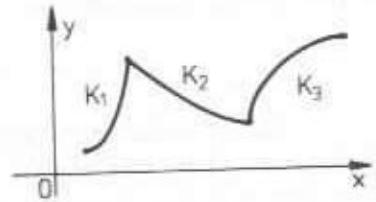
$$\int_K f(x,y)dl.$$

Zatem

$$\int_K f(x,y)dl \overset{\text{def}}{=} \lim_{n\to\infty}\sum_{i=1}^{n} f(\bar{x}_i,\bar{y}_i)\Delta l_i.$$

Załóżmy, że krzywą kawałkami gładką K można podzielić (por. rys. 6.2) na łuki gładkie $K_1,K_2,\ldots,K_n$. Całkę krzywoliniową nieskierowaną funkcji f po krzywej K określamy równością

$$\int_K f(x,y)dl \overset{\text{def}}{=} \sum_{i=1}^{n}\int_{K_i} f(x,y)dl.$$



Rys 6.2

TWIERDZENIE 6.1 (warunek wystarczający istnienia całki krzywoliniowej nieskierowanej). Jeżeli f jest funkcją ciągłą na łuku gładkim K, to całka krzywoliniowa $\int_K f(x,y)dl$ istnieje.

INTERPRETACJA FIZYCZNA CAŁKI KRZYWOLINIOWEJ NIESKIEROWANEJ. Masa m jednorodnego łuku gładkiego o długości l i gęstości ρ jest równa

$$m=\rho\cdot l.$$

Zdefiniujemy teraz masę łuku gładkiego niejednorodnego, którego gęstość liniowa w punkcie (x,y) jest równa $\rho(x,y)$, przy czym ρ jest funkcją ciągłą (i oczywiście nieujemną) na łuku K.

Podzielmy łuk K jak w definicji całki krzywoliniowej nieskierowanej. Iloczyn $\rho(\overline{x}_i,\overline{y}_i)\Delta l_i$ jest równy masie jednorodnego łuku częściowego o długości Δl_i oraz stałej gęstości $\rho(\overline{x}_i,\overline{y}_i)$ i jest przybliżoną masą łuku częściowego niejednorodnego. Zatem suma

$$\sum_{i=1}^{n}\rho(\overline{x}_i,\overline{y}_i)\Delta l_i$$

jest wartością przybliżoną masy niejednorodnego łuku K. Rozważmy teraz granicę

$$\lim_{n\to\infty}\sum_{i=1}^{n}\rho(\overline{x}_i,\overline{y}_i)\Delta l_i .$$

Ponieważ funkcja ρ jest ciągła na łuku gładkim, więc granica ta istnieje i jest równa całce $\int_K \rho(x,y)\,dl$. Naturalnym jest więc przyjąć następującą definicję masy łuku gładkiego

$$m \overset{\text{def}}{=} \lim_{n\to\infty}\sum_{i=1}^{n}\rho(\overline{x}_i,\overline{y}_i)\Delta l_i .$$

Zatem: całka krzywoliniowa nieskierowana ciągłej i nieujemnej funkcji $\rho(x,y)$ po łuku gładkim K jest równa masie m łuku K o gęstości $\rho(x,y)$ w każdym punkcie (x,y) tego łuku:

$$\int_K \rho(x,y)\,dl = m. \tag{6.2}$$

ZAMIANA CAŁKI KRZYWOLINIOWEJ NIESKIEROWANEJ NA CAŁKĘ OZNACZONĄ. Obliczanie całki krzywoliniowej nieskierowanej z funkcji ciągłej sprowadza się do obliczania odpowiedniej całki oznaczonej. Wynika to z następującego twierdzenia.

TWIERDZENIE 6.2. Jeżeli funkcja f jest ciągła na łuku gładkim

$$K:\ x=x(t),\ y=y(t),\quad t\in\langle\alpha,\beta\rangle,$$

to

$$\int_K f(x,y)\,dl = \int_\alpha^\beta f(x(t),y(t))\sqrt{[x'(t)]^2+[y'(t)]^2}\,dt. \tag{6.3}$$

W szczególności, gdy łuk gładki K określony jest równaniem jawnym:

$$y=y(x),\quad x\in\langle a,b\rangle$$

wzór (6.3) przyjmuje postać

$$\int_K f(x,y)\,dl = \int_a^b f(x,y(x))\sqrt{1+[y'(x)]^2}\,dx .$$

Istotnie, łuk gładki K określony równaniem $y=y(x)$, $x\in\langle a,b\rangle$ można zapisać równaniami parametrycznymi

$$x=t,\ y=y(t),\quad t\in\langle a,b\rangle,$$

a po zastosowaniu twierdzenia 6.2 otrzymujemy powyższy wzór (wprawdzie zmienna pod całką po prawej stronie będzie oznaczona literą t, ale jest to zmienna pozorna).

PRZYKŁAD 6.1. Obliczymy następujące całki krzywoliniowe nieskierowane:

a) $\int_K (y-1)\,dl$, jeśli K jest łukiem paraboli $x=t^2+1$, $y=t+1$, $t\in\langle 0,\sqrt{2}\rangle$,

b) $\int_K x^2\,dl$, jeśli K jest okręgiem o równaniu $x^2+y^2=9$,

c) $\int_K \frac{x^2}{\sqrt{x^2+16y^2}}\,dl$, jeśli K jest elipsą $x^2+4y^2=4$.

Każdą z tych całek obliczymy korzystając z twierdzenia 6.2, przy czym w przykładach b) i c) napiszemy najpierw równania parametryczne krzywej całkowania.

a) $\int_K (y-1)\,dl = \left\{\begin{matrix} K:\ x=t^2+1,\ y=t+1,\ 0\le t\le\sqrt{2} \\ (x'(t))^2+(y'(t))^2=4t^2+1 \end{matrix}\right\} = \int_0^{\sqrt{2}} t\sqrt{4t^2+1}\,dt =$

$$= \frac{1}{12}\left[(4t^2+1)^{\frac{3}{2}}\right]_0^{\sqrt{2}} = \frac{13}{6}.$$

b) Okrąg $x^2+y^2=9$ opisujemy równaniami parametrycznymi

$$x=3\cos t,\quad y=3\sin t,\quad t\in\langle 0,2\pi\rangle,$$

a następnie korzystając z twierdzenia 6.2 obliczamy:

$$\int_K x^2\,dl = \left\{\begin{array}{l}K:\ x=3\cos t,\ y=3\sin t,\ t\in\langle 0,2\pi\rangle \\ (x'(t))^2+(y'(t))^2=(-3\sin t)^2+(3\cos t)^2=9\end{array}\right\} = \int_0^{2\pi} 27\cos^2 t\,dt =$$

$$= \frac{27}{2}\left[t+\frac{1}{2}\sin 2t\right]_0^{2\pi} = 27\pi.$$

c) Elipsę $K: x^2/4+y^2/1=1$ opisujemy równaniami parametrycznymi

$$x=2\cos t,\quad y=\sin t,\quad t\in\langle 0,2\pi\rangle,$$

i obliczamy:

$$\int_K \frac{x^2}{\sqrt{x^2+16y^2}}\,dl = \left\{\begin{array}{l}K:\ x=2\cos t,\ y=\sin t,\ t\in\langle 0,2\pi\rangle \\ (x'(t))^2+(y'(t))^2=4\sin^2 t+\cos^2 t\end{array}\right\} =$$

$$= \int_0^{2\pi} \frac{4\cos^2 t}{\sqrt{4\cos^2 t+16\sin^2 t}}\sqrt{4\sin^2 t+\cos^2 t}\,dt = \int_0^{2\pi} 2\cos^2 t\,dt =$$

$$= \left[t+\frac{1}{2}\sin 2t\right]_0^{2\pi} = 2\pi. \quad \blacksquare$$

PRZYKŁAD 6.2. Obliczymy masę m odcinka K o końcach A(1,3), B(2,6), którego gęstość liniowa jest równa $\rho(x,y)=xy$.

Zgodnie ze wzorem (6.2) mamy

$$m = \int_K xy\,dl = \left\{\begin{array}{l}K:\ x=t,\ y=3t,\ t\in\langle 1,2\rangle \\ (x'(t))^2+(y(t))^2=10\end{array}\right\} = \int_1^2 3t^2\sqrt{10}\,dt = \sqrt{10}\,t^3\Big|_1^2 = 7\sqrt{10}.$$

$\blacksquare$

Analogicznie określa się i oblicza całkę krzywoliniową nieskierowaną funkcji $f(x,y,z)$ po łuku gładkim K w przestrzeni R^3.

Jeśli funkcja $f(x,y,z)$ jest ciągła na łuku gładkim

$$K:\ x=x(t),\ y=y(t),\ z=z(t),\quad t\in\langle\alpha,\beta\rangle,$$

to

$$\int_K f(x,y,z)\,dl = \int_\alpha^\beta f(x(t),y(t),z(t))\sqrt{(x'(t))^2+(y'(t))^2+(z'(t))^2}\,dt.$$

PRZYKŁAD 6.3. Obliczymy całkę $\int_K \frac{x}{z}\,dl$ po łuku K określonym równaniami parametrycznymi $x=t,\ y=\sqrt{3}t,\ z=1+t^2,\ t\in\langle 0,1\rangle$.

$$\int_K \frac{x}{z}\,dl = \left\{\begin{array}{l}K:\ x=t,\ y=\sqrt{3}t,\ z=1+t^2,\ t\in\langle 0,1\rangle \\ (x'(t))^2+(y'(t))^2+(z'(t))^2=4+4t^2\end{array}\right\} = \int_0^1 \frac{2t\sqrt{1+t^2}}{1+t^2}\,dt =$$

$$= \int_0^1 \frac{2t}{\sqrt{1+t^2}}\,dt = \left[2\sqrt{1+t^2}\right]_0^1 = 2\sqrt{2}-2. \quad \blacksquare$$

ZADANIA DO ROZWIĄZANIA.

1. Obliczyć całki krzywoliniowe nieskierowane:

a) $\int_K xy^2\,dl$, jeśli $K: x=-1+t, y=1+t, t\in\langle 0,2\rangle$,

b) $\int_K \frac{2x+1}{y}\,dl$, jeśli K jest odcinkiem o końcach A(-2,1), B(0,3),

c) $\int_K y\sqrt{3+x}\,dl$, jeśli K jest odcinkiem o końcach A(-2,2), B(1,-1),

d) $\int_K \frac{y}{x+2}\,dl$, jeśli K jest łamaną ABC o wierzchołkach A(-1,0),

B(1,2), C(1,-1),

e) $\int_K (x+1)y\,dl$, jeśli K: $x=t^2$, $y=2t$, $t\in\langle 0,\sqrt{3}\rangle$,

f) $\int_K xy\,dl$, jeśli K: $y=\frac{1}{2}x^2$, $x\in\langle 0,\sqrt{3}\rangle$,

g) $\int_K \frac{x}{y+1}\,dl$, jeśli K: $2y-x^2+1=0$, $x\in\langle 0,2\rangle$,

h) $\int_K \frac{\sqrt{x+1}}{y}\,dl$, jeśli K: $y=2\sqrt{x}$, $x\in\langle 2,4\rangle$,

i) $\int_K \frac{x^2y}{\sqrt{1+x^2}}\,dl$, jeśli K: $y=\ln x$, $x\in\langle 1,e\rangle$,

j) $\int_K \frac{y}{4+x}\,dl$, jeśli K: $x=\cos t$, $y=\sin t$, $t\in\langle 0,\pi\rangle$,

k) $\int_K (x^2-y^2)\,dl$, jeśli K: $x=\cos t$, $y=\sin t$, $t\in\langle 0,\pi/4\rangle$,

l) $\int_K (y^2-xy)\,dl$, jeśli K: $x=\cos t$, $y=\sin t$, $t\in\langle 0,\pi\rangle$,

ł) $\int_K x^2y^2\,dl$, jeśli K: $x^2+y^2=4$,

m) $\int_K (yx^2+1)\,dl$, jeśli K: $x^2+y^2=4$, $y\geq 0$,

n) $\int_K x^3\,dl$, jeśli K: $x^2+y^2=4$, $x\geq 0$,

o) $\int_K xy^2\,dl$, jeśli K: $x^2+y^2-2x-3=0$,

p) $\int_K y^2\,dl$, jeśli K: $x^2+2x+y^2-6y+6=0$,

r) $\int_K xy\,dl$, jeśli K: $x=3\cos t$, $y=2\sin t$, $t\in\langle 0,\pi/2\rangle$,

s) $\int_K y\,dl$, jeśli K: $x=\cos t+t\sin t$, $y=\sin t-t\cos t$, $t\in\langle 0,\pi\rangle$,

t) $\int_K y\,dl$, jeśli K: $x=\cos^3 t$, $y=\sin^3 t$, $t\in\langle 0,\pi/2\rangle$,

u) $\int_K x\,dl$, jeśli K: $x=\cos^3 t$, $y=\sin^3 t$, $t\in\langle -\pi/2,\pi/2\rangle$,

w) $\int_K \frac{y}{1+2z}\,dl$, jeśli K: $x=t$, $y=t$, $z=t^2$, $t\in\langle 0,2\rangle$.

2. Obliczyć masę m łuku K, którego gęstość liniowa jest równa $\rho(x,y)$, jeśli:

a) K jest okręgiem ośrodku (2,0) i promieniu 3, $\rho(x,y)=x^2$,

b) K jest łukiem paraboli $y^2=x$ dla $y\in\langle -1,1\rangle$, $\rho(x,y)=\sqrt{x}$,

c) K jest łamaną ABC, przy czym A(0,0), B(2,2), C(0,4), $\rho(x,y)=x$,

d) K jest odcinkiem $\overline{AB}$, A(2,1,0), B(4,2,1), $\rho(x,y)=x+y$.

Odpowiedzi.

1. a) $8\sqrt{2}/3$, b) $\sqrt{2}(4-5\ln 3)$, c) $8\sqrt{2}/5$, d) $\sqrt{2}(2-\ln 3)+1/2$, e) 124/5, f) 29/15, g) $2(\sqrt{5}-1)$, h) $1+\ln\sqrt{2}$, i) $(e^2+1)/4$, j) $\ln 5-\ln 3$, k) 1/2, l) $\pi/2$, ł) 8π, m) $2\pi+32/3$, n) 64/3, o) 8π, p) 44π, r) 38/5, s) 3π, t) 3/5, u) 6/5, w) $\sqrt{2}$.

2. a) $m=51\pi$, L: $x=2+3\cos t$, $y=3\sin t$, $t\in\langle 0,2\pi\rangle$,
b) $m=(5\sqrt{5}-1)/6$, c) $m=4\sqrt{2}$,
d) $9\sqrt{6}/2$; $x=2+2t$, $y=1+t$, $z=t$, $0\leq t\leq 1$.

7. CAŁKA KRZYWOLINIOWA SKIEROWANA.

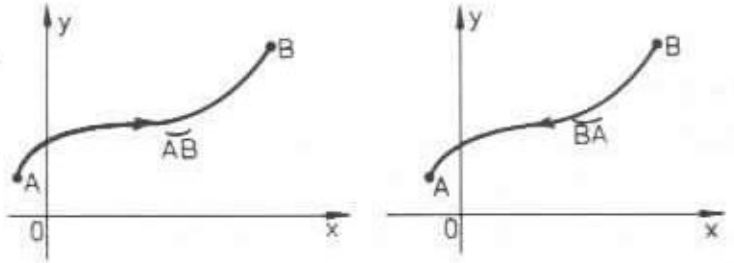
ŁUK GŁADKI SKIEROWANY. W dalszym ciągu będzie nam potrzebne pojęcie łuku gładkiego s k i e r o w a n e g o. Przyjmijmy następujące określenie.

Łuk gładki, w którym jeden z dwóch końców wyróżniono nazywając go początkiem, a drugi - końcem, nazywamy **łukiem gładkim skierowanym (zorientowanym)**.

Niech będzie dany łuk gładki określony równaniami parametrycznymi

(7.1) $$x=x(t),\quad y=y(t),\quad t\in\langle\alpha,\beta\rangle.$$

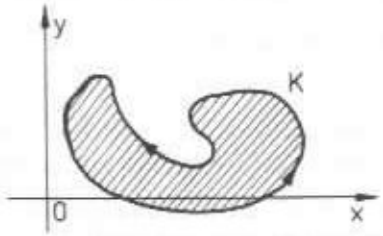
Łuk skierowany opisany tymi równaniami, którego początkiem jest punkt $A=(x(\alpha),y(\alpha))$, a końcem punkt $B=(x(\beta),y(\beta))$ oznaczmy symbolem $\overset{\frown}{AB}$. Mówimy, że skierowanie łuku $\overset{\frown}{AB}$ jest zgodne z jego parametryzacją. Łuk skierowany opisany równaniami (7.1), którego początkiem jest punkt B, a końcem punkt A oznaczmy symbolem $\overset{\frown}{BA}$. Łuki $\overset{\frown}{AB}$ i $\overset{\frown}{BA}$ nazywamy łukami przeciwnie skierowanymi i zapisujemy $\overset{\frown}{BA}=-\overset{\frown}{AB}$ (por. rys. 7.1). Łuk przeciwnie skierowany do łuku K oznaczamy przez $-K$.



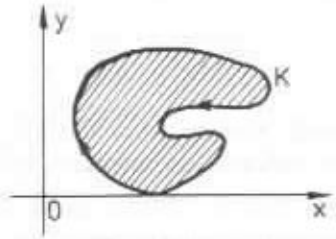
Rys 7.1

Dla określenia łuku skierowanego są więc potrzebne nie tylko równania (równanie), ale także wskazanie początku i końca tego łuku. Jeżeli jednak w dalszym ciągu naszych rozważań podamy jedynie równania (równanie) łuku skierowanego, to będziemy przyjmować, że skierowanie tego łuku jest „naturalne", tzn. zgodne z jego parametryzacją.

Mówimy, że krzywa zamknięta kawałkami gładka jest **skierowana dodatnio**, gdy obserwator poruszający się po tej krzywej zgodnie z jej skierowaniem ma „najbliższe" punkty wnętrza tej krzywej po swojej lewej ręce (rys. 7.2). W przeciwnym razie mówimy, że krzywa jest **skierowana ujemnie** (rys. 7.3).

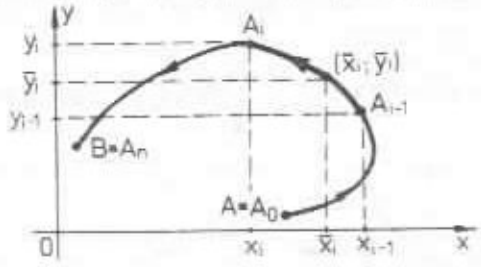


Rys 7.2



Rys 7.3

OKREŚLENIE CAŁKI KRZYWOLINIOWEJ SKIEROWANEJ. Rozważmy dwie funkcje P i Q określone w każdym punkcie (x,y) łuku gładkiego skierowanego K o początku A i końcu B. Analogicznie, jak w definicji całki krzywoliniowej nieskierowanej, podzielmy łuk K punktami $A=A_0,A_1,A_2,\ldots,A_{n-1},A_n=B$ na n łuków częściowych $l_1,l_2,\ldots,l_n$, gdzie $l_i=\overset{\frown}{A_{i-1}A_i}$, $i=1,2,\ldots,n$. (rys 7.4). Niech Δl_i oznacza długość łuku l_i, $i=1,2,\ldots,n$. Liczbę $\delta_n=\max(\Delta l_1,\Delta l_2,\ldots,\Delta l_n)$ nazywamy średnicą danego podziału. Załóżmy, że $A_i=(x_i,y_i)$ i przyjmijmy oznaczenia: $\Delta x_i=x_i-x_{i-1}$, $\Delta y_i=y_i-y_{i-1}$, $i=1,2,\ldots,n$.



Rys 7.4

Na każdym łuku częściowym l_i wybierzmy punkt pośredni $(\bar{x}_i,\bar{y}_i)$ i utwórzmy sumę

$$S_n=\sum_{i=1}^{n}[P(\bar{x}_i,\bar{y}_i)\Delta x_i+Q(\bar{x}_i,\bar{y}_i)\Delta y_i].$$

Utwórzmy następnie normalny ciąg podziałów łuku K na łuki częściowe, (tzn., że $\delta_n\to 0$ przy $n\to\infty$) i odpowiadający mu ciąg (S_n). Rozważmy granicę

$$\lim_{n\to\infty} S_n=\lim_{n\to\infty}\sum_{i=1}^{n}[P(\bar{x}_i,\bar{y}_i)\Delta x_i+Q(\bar{x}_i,\bar{y}_i)\Delta y_i].$$

Jeżeli dla każdego normalnego ciągu podziałów łuku K i każdego wyboru punktów pośrednich $(\bar{x}_i,\bar{y}_i)$ na łukach częściowych istnieje ta sama skończona granica ciągu (S_n), to tę granicę nazywamy **całką krzywoliniową skierowaną (zorientowaną)** pary funkcji P, Q po łuku gładkim skierowanym K o początku A i końcu B i oznaczamy symbolem

$$\int_K P(x,y)dx+Q(x,y)dy \quad \text{lub} \quad \int_{AB} P(x,y)dx+Q(x,y)dy.$$

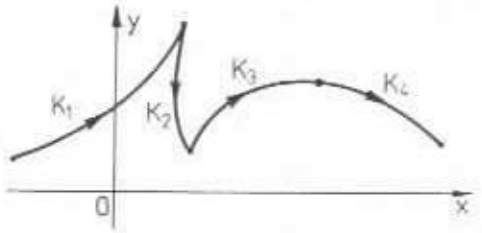
Zatem

$$\int_K P(x,y)dx+Q(x,y)dy \overset{def}{=} \lim_{n\to\infty}\sum_{i=1}^{n}[P(\bar{x}_i,\bar{y}_i)\Delta x_i+Q(\bar{x}_i,\bar{y}_i)\Delta y_i].$$

Z definicji tej wynika, że:

Jeżeli istnieje całka krzywoliniowa skierowana pary funkcji P i Q po łuku skierowanym K, to istnieje całka tych funkcji po łuku przeciwnie skierowanym $-K$, przy czym

$$\int_{-K} P(x,y)dx+Q(x,y)dy=-\int_K P(x,y)dx+Q(x,y)dy$$



Rys 7.5

Załóżmy, że K jest krzywą skierowaną kawałkami gładką, którą można podzielić (por. rys. 7.5) na n łuków gładkich $K_1,K_2,\ldots,K_n$. Całkę krzywoliniową skierowaną funkcji P i Q po krzywej K określamy równością

$$(7.2)\qquad \int_K P(x,y)dx+Q(x,y)dy \overset{def}{=} \sum_{i=1}^{n}\int_{K_i} P(x,y)dx+Q(x,y)dy.$$

Całkę krzywoliniową skierowaną funkcji P i Q po krzywej z a m k n i ę t e j K oznacza się także symbolem

$$\oint_K P(x,y)dx+Q(x,y)dy.$$

TWIERDZENIE 7.1 Ciągłość funkcji P i Q na łuku gładkim skierowanym K jest warunkiem wystarczającym istnienia całki

$$\int_K P(x,y)dx+Q(x,y)dy.$$

Inaczej: Jeżeli funkcje P i Q są ciągłe na łuku gładkim skierowanym K, to są całkowalne na tym łuku.

INTERPRETACJA FIZYCZNA CAŁKI KRZYWOLINIOWEJ SKIEROWANEJ. Praca W s t a ł e j siły $\bar{F}$ na drodze prostoliniowej $\bar{s}$ jest równa iloczynowi skalarnemu:

$$W=\bar{F}\circ\bar{s}.$$

Oznaczając przez P, Q oraz $\Delta x,\Delta y$ współrzędne odpowiednio siły $\bar{F}$ i drogi $\bar{s}$ otrzymujemy

$$W=P\Delta x+Q\Delta y.$$

Zdefiniujemy teraz pracę z m i e n n e j siły

$$\bar{F}(x,y)=[P(x,y),Q(x,y)]$$

na łuku gładkim skierowanym $K\subset 0xy$, przy założeniu, że funkcje P, Q są ciągłe na tym łuku.

Podzielmy łuk K tak, jak w definicji całki krzywoliniowej skierowanej. Iloczyn skalarny $P(\bar{x}_i,\bar{y}_i)\Delta x_i+Q(\bar{x}_i,\bar{y}_i)\Delta y_i$ jest równy pracy stałej siły $\bar{F}=[P(\bar{x}_i,\bar{y}_i),Q(\bar{x}_i,\bar{y}_i)]$ na drodze prostoliniowej

$A_{i-1}A_i=[\Delta x_i,\Delta y_i]$ i jest równy w przybliżeniu pracy zmiennej siły $\bar{F}(x,y)$ na łuku skierowanym $A_{i-1}A_i$. Zatem suma

$$\sum_{i=1}^{n}[P(\bar{x}_i,\bar{y}_i)\Delta x_i+Q(\bar{x}_i,\bar{y}_i)\Delta y_i]$$

jest wartością przybliżoną pracy zmiennej siły $[P(x,y),Q(x,y)]$ na drodze krzywoliniowej K.

Naturalnym jest więc przyjąć następującą definicję pracy zmiennej siły $\bar{F}(x,y)=[P(x,y),Q(x,y)]$ na łuku gładkim skierowanym K:

$$W=\lim_{n\to\infty}\sum_{i=1}^{n}[P(\bar{x}_i,\bar{y}_i)\Delta x_i+Q(\bar{x}_i,\bar{y}_i)\Delta y_i].$$

Ponieważ funkcje P, Q są ciągłe na łuku K, więc granica ta istnieje i jest równa całce krzywoliniowej skierowanej. Zatem:

Całka krzywoliniowa skierowana pary funkcji P, Q ciągłych na gładkim łuku skierowanym K jest równa pracy zmiennej siły $\bar{F}(x,y)=[P(x,y),Q(x,y)]$ na łuku K:

$$\int_K P(x,y)dx+Q(x,y)dy=W. \tag{7.3}$$

ZAMIANA CAŁKI KRZYWOLINIOWEJ SKIEROWANEJ NA OZNACZONĄ. Obliczanie całki krzywoliniowej skierowanej po łuku gładkim z funkcji ciągłych sprowadza się do obliczenia odpowiedniej całki oznaczonej, co wynika z następującego twierdzenia.

TWIERDZENIE 7.2. Jeżeli funkcje P, Q są ciągłe na łuku gładkim

$$K:\ x=x(t),\ y=y(t),\quad t\in<\alpha,\beta>, \tag{7.4}$$

skierowanym zgodnie z jego parametryzacją, to

$$\int_K P(x,y)dx+Q(x,y)dy=\int_\alpha^\beta[P(x(t),y(t))x'(t)+Q(x(t),y(t))y'(t)]dt. \tag{7.5}$$

Uwaga: Jeżeli łuk K jest określony równaniami (7.4), ale jego początkiem jest punkt $(x(\beta),y(\beta))$, a końcem punkt $(x(\alpha),y(\alpha))$, to

$$\int_K P(x,y)dx+Q(x,y)dy=\int_\beta^\alpha[P(x(t),y(t))x'(t)+Q(x(t),y(t))y'(t)]dt \tag{7.6}$$

Z twierdzenia 7.2 wynika, że:

1) Jeśli łuk skierowany K jest określony równaniem

$$y=y(x),\ x\in<a,b>,$$

gdzie $y(x)$ jest funkcją klasy C^1 na przedziale $<a,b>$, to wzór (7.5) przyjmuje postać

$$\int_K P(x,y)dx+Q(x,y)dy=\int_a^b[P(x,y(x))+Q(x,y(x))y'(x)]dx. \tag{7.7}$$

2) Jeśli łuk skierowany K jest odcinkiem prostopadłym do osi 0y

$$K:\ y=c,\quad x\in<a,b>,$$

to

$$\int_K P(x,y)dx+Q(x,y)dy=\int_a^b P(x,c)dx.$$

3) Jeśli łuk skierowany K jest odcinkiem prostopadłym do osi 0x

$$K:\ x=a,\quad y\in<c,d>,$$

to

$$\int_K P(x,y)dx+Q(x,y)dy=\int_c^d Q(a,y)dy.$$

PRZYKŁAD 7.1. Obliczymy następujące całki krzywoliniowe:

a) $\int_K xdx+xydy$, jeśli $K:\ x=2\cos t,\ y=2\sin t,\ t\in<0,\pi>$,

b) $\int_K \sqrt{y-x}dx+xdy$, jeśli K jest odcinkiem o początku w punkcie (2,3) i końcu w punkcie (0,9),

c) $\int_K (2-\frac{6x}{y+1})dx+\frac{1}{x}\ln x\,dy$, jeśli K oznacza łuk paraboli

$y=x^2-1$ od punktu (2,3) do punktu (1,0).

d) $\int_K y\cos(y-x)dx$, jeśli K jest łamaną ABC o wierzchołkach A(0,0), B(2,2), C(2,-1).

W każdym z tych przykładów obliczamy zadaną całkę krzywoliniową zamieniając ją na całkę oznaczoną.

a) Zgodnie z twierdzeniem 7.2 mamy:

$$\int_K xdx+xydy=\left\{\begin{matrix}K:\ x=2\cos t,\ y=2\sin t,\ t\in\langle 0,\pi\rangle\\ x'(t)=-2\sin t,\ y'(t)=2\cos t\end{matrix}\right\}=$$

$$=\int_0^{\pi}(2\cos t(-2\sin t)+4\cos t\sin t\cdot 2\cos t)dt=\int_0^{\pi}(-2\sin 2t+8\cos^2 t\sin t)dt=$$

$$=\left[\cos 2t-\frac{8}{3}\cos^3 t\right]_0^{\pi}=\frac{16}{3}.$$

b) Przypomnijmy, że równania parametryczne odcinka o początku w punkcie (x_1,y_1) i końcu w punkcie (x_2,y_2) mają postać

$$x=x_1+(x_2-x_1)t,\quad y=y_1+(y_2-y_1)t,\quad t\in\langle 0,1\rangle$$

(dla $t=0$ otrzymujemy początek, a dla $t=1$ koniec tego odcinka).

Obliczamy

$$\int_K\sqrt{y-x}dx+xdy=\left\{\begin{matrix}K:\ x=2-2t,\ y=3+6t,\ t\in\langle 0,1\rangle\\ x'(t)=-2,\ y'(t)=6\end{matrix}\right\}=$$

$$=\int_0^1(-2\sqrt{1+8t}+6(2-2t))dt=\left[-\frac{1}{6}(\sqrt{1+8t})^3+12t-6t^2\right]_0^1=\frac{5}{3}.$$

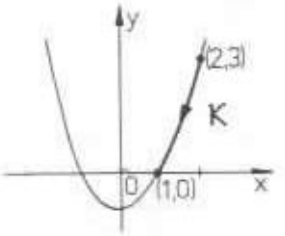
c) Krzywą całkowania K (rys 7.6) można opisać równaniami parametrycznymi

$$x=t,\quad y=t^2-1,\quad t\in\langle 1,2\rangle,$$

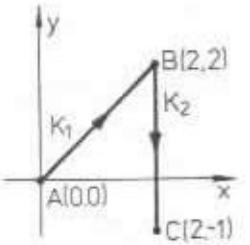
przy czym początek tej krzywej otrzymujemy dla $t=2$, a koniec dla $t=1$. Zatem zgodnie ze wzorem (7.6) mamy

$$\int_K(2-\frac{6x}{y+1})dx+\frac{1}{x}\ln x dy=\int_2^1(2-\frac{6t}{t^2}+\frac{1}{t}\ln t\cdot 2t)dt=\int_2^1(2-\frac{6}{t}+2\ln t)dt=$$

$$=(2t-6\ln t+2t\ln t-2t)\Big|_2^1=\ln 4.$$



Rys 7.6



Rys 7.7

d) W przykładzie tym $Q(x,y)\equiv 0$. Krzywą całkowania widzimy na rysunku 7.7. Zatem, zgodnie z (7.2) i (7.5), mamy

$$\int_K y\cos(y-x)dx=\int_{K_1}+\int_{K_2}=\left\{\begin{matrix}K_1:\ x=2t,\ y=2t,\ t\in\langle 0,1\rangle\\ K_2:\ x=2,\ y=2-3t,\ t\in\langle 0,1\rangle\end{matrix}\right\}=$$

$$=\int_0^1 4t\cos 0\,dt+\int_0^1(2-3t)\cos(-3t)\cdot 0\,dt=2t^2\Big|_0^1=2.\quad\blacksquare$$

PRZYKŁAD 7.2. Obliczymy pracę W zmiennej siły $\vec{F}=[x,x]$ na łuku K paraboli $y=x^2$ od punktu A(-1,1) do punktu B(3,9).

Zgodnie ze wzorem (7.3) mamy

$$W=\int_K xdx+xdy=\left\{\begin{matrix}K:\ x=t,\ y=t^2,\ t\in\langle -1,3\rangle\\ x'(t)=1,\ y'(t)=2t\end{matrix}\right\}=$$

$$=\int_{-1}^3(t+2t^2)dt=\left[\frac{1}{2}t^2+\frac{2}{3}t^3\right]_{-1}^3=\frac{68}{3}.\quad\blacksquare$$

Analogicznie definiujemy i obliczamy całkę krzywoliniową skierowaną funkcji P(x,y,z), Q(x,y,z), R(x,y,z) po krzywej w przestrzeni R^3.

Jeżeli funkcje P, Q, R są ciągłe na łuku gładkim

$$K:\ x=x(t),\ y=y(t),\ z=z(t),\ t\in\langle\alpha,\beta\rangle,$$

skierowanym zgodnie z jego parametryzacją, to

$$\int_K P(x,y,z)dx+Q(x,y,z)dy+R(x,y,z)dz=$$

$$=\int_\alpha^\beta [P(x(t),y(t),z(t))x'(t)+Q(x(t),y(t),z(t))y'(t)+R(x(t),y(t),z(t))z'(t)]dt.$$

Obliczymy, na przykład, całkę

$$\int_K (x^2-4z)dx+(x-z^2)dy+(z^2-y^2)dz,$$

gdzie K oznacza odcinek o początku (1,2,1) i końcu (3,1,2).

Piszemy równania parametryczne odcinka K:

$$x=1+2t,\quad y=2-t,\quad z=1+t,\quad t\in\langle 0,1\rangle,$$

i obliczamy

$$\int_K (x^2-4z)dx+(x-z^2)dy+(z^2-y^2)dz=\int_0^1 \left((4t^2-3)\cdot 2+t^2+6t-3\right)dt=$$

$$=\int_0^1 (9t^2+6t-9)dt=(3t^3-3t^2-9t)\Big|_0^1=-9.$$

ZWIĄZEK CAŁKI KRZYWOLINIOWEJ SKIEROWANEJ Z CAŁKĄ PODWÓJNĄ. Obliczanie całki krzywoliniowej skierowanej po krzywej zamkniętej można, jeżeli tylko spełnione są odpowiednie założenia, sprowadzić do obliczenia całki podwójnej po obszarze ograniczonym tą krzywą. Wynika to z następującego twierdzenia.

TWIERDZENIE GREENA. Jeżeli funkcje $P(x,y)$, $Q(x,y)$ oraz ich pochodne P'_y, Q'_x są ciągłe na obszarze D normalnym względem obu osi układu współrzędnych, a brzegiem obszaru D jest skierowana dodatnio krzywa kawałkami gładka K, to

$$\int_K P(x,y)dx+Q(x,y)dy=\iint_D (Q'_x(x,y)-P'_y(x,y))dxdy.$$

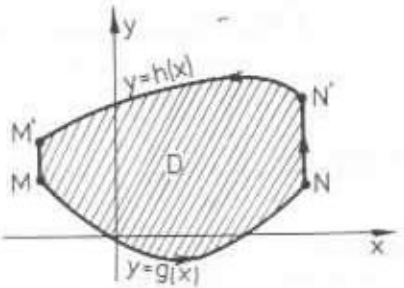
D o w ó d . Wystarczy udowodnić, że

(1) $$\int_K P(x,y)dx=-\iint_D P'_y(x,y)dxdy,$$

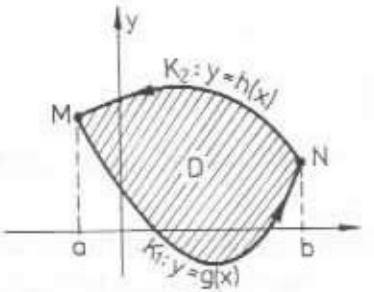
(2) $$\int_K Q(x,y)dy=\iint_D Q'_x(x,y)dxdy.$$

Obszar D jest normalny względem osi 0x (rys. 7.8 - 7.9), zatem

$$D=\{(x,y)\in R^2:\ a\le x\le b\ \wedge\ g(x)\le y\le h(x)\}.$$



Rys 7.8



Rys 7.9

Dla uproszczenia załóżmy, że brzeg K obszaru D (tak, jak na rysunku 7.9) nie zawiera odcinków prostych $x=a$, $x=b$.

Podzielmy krzywą K punktami M i N na krzywe K_1 i K_2, przy czym

K_1 oznacza krzywą o równaniu $y=g(x)$, $x\in\langle a,b\rangle$, o początku w punkcie M i końcu w punkcie N (początek tej krzywej otrzymujemy dla $x=a$, a koniec dla $x=b$),

K_2 oznacza krzywą o równaniu $y=h(x)$, $x\in\langle a,b\rangle$, o początku w punkcie N i końcu w punkcie M (początek tej krzywej otrzymujemy dla $x=b$, a koniec dla $x=a$).

Zatem, zgodnie z (7.2) i (7.7), mamy

$$\int_K P(x,y)dx=\int_{K_1} P(x,y)dx+\int_{K_2} P(x,y)dx=\int_a^b P(x,g(x))dx+\int_b^a P(x,h(x))dx=$$

$$=\int_a^b P(x,g(x))dx-\int_a^b P(x,h(x))dx,$$

czyli

$$\text{(a)}\quad \int_K P(x,y)dx = \int_a^b (P(x,g(x)) - P(x,h(x))dx .$$

Następnie weźmy pod uwagę prawą stronę równości (1):

$$-\iint_D P'_y(x,y)dxdy = -\int_a^b [\int_{g(x)}^{h(x)} P'_y(x,y)dy]dx = -\int_a^b P(x,y)\Big|_{g(x)}^{h(x)} dx =$$

$$= -\int_a^b (P(x,h(x)) - P(x,g(x))dx .$$

Stąd

$$\text{(b)}\quad -\iint_D P'_y(x,y)dxdy = \int_a^b (P(x,g(x)) - P(x,h(x))dx .$$

Z (a) i (b) wynika natychmiast równość (1).

Analogicznie, wykorzystując założenie normalności obszaru D względem osi 0y, uzyskujemy równość (2), co kończy dowód. □

Uwaga 1. Zauważmy, że dowód dla obszaru przedstawionego na rysunku 7.8 nie zmienia się istotnie, gdyż $\int_K P(x,y)dx$ po odcinkach skierowanych NN' i M'M jest równa zeru.

Uwaga 2. Twierdzenie Greena pozostaje prawdziwe dla dowolnego obszaru, który można podzielić na skończoną liczbę obszarów normalnych względem obu osi układu i którego brzegiem jest skierowana dodatnio krzywa kawałkami gładka.

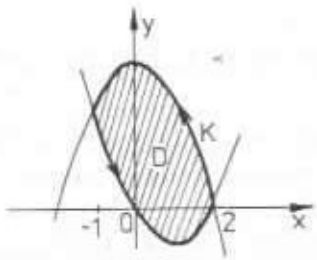
PRZYKŁAD 7.3. Obliczymy następujące całki krzywoliniowe:

a) $\int_K (xy + e^{-x^2})dx + (x^2 - e^{-y^2})dy$, jeśli K jest skierowanym dodatnio brzegiem obszaru D ograniczonego parabolami $y = 4 - x^2$, $y = x^2 - 2x$,

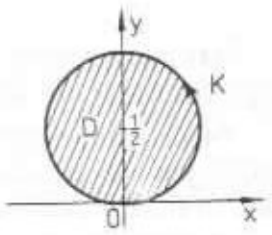
b) $\int_K (2y + \frac{x}{1+x^2+y^2})dx + (3x^2 + \frac{y}{1+x^2+y^2})dy$, jeśli K jest dodatnio skierowanym okręgiem o równaniu $x^2 + y^2 - y = 0$.

a) Krzywa całkowania K jest skierowanym dodatnio brzegiem obszaru (rys. 7.10)

$$D = \{(x,y) \in R^2 : -1 \le x \le 2 \wedge x^2 - 2x \le y \le 4 - x^2\} .$$



Rys 7.10



Rys 7.11

Stosując twierdzenie Greena otrzymujemy

$$\int_K (xy + e^{-x^2})dx + (x^2 - e^{-y^2})dy = \begin{Bmatrix} Q'_x = 2x,\ P'_y = x \\ Q'_x - P'_y = x \end{Bmatrix} = \iint_D x dxdy =$$

$$= \int_{-1}^{2} [\int_{x^2-2x}^{4-x^2} x dy]dx = \int_{-1}^{2} x(4 + 2x - 2x^2)dx = \left[2x^2 + \frac{2}{3}x^3 - \frac{1}{2}x^4\right]_{-1}^{2} = \frac{9}{2} .$$

b) Korzystając z twierdzenia Greena zamienimy daną całkę krzywoliniową na całkę podwójną po kole D określonym nierównością $x^2 + y^2 - y \le 0$, (rys. 7.11). Ponieważ

$$Q'_x = 6x - \frac{2xy}{(1+x^2+y^2)^2},\quad P'_y = 2 - \frac{2xy}{(1+x^2+y^2)^2},\quad Q'_x - P'_y = 6x - 2,$$

zatem

$$\int_K (2y + \frac{x}{1+x^2+y^2})dx + (3x^2 + \frac{y}{1+x^2+y^2})dy = \iint_D (6x - 2)dxdy =$$

$$= \left\{ \begin{array}{l} x = r\cos\varphi,\ y = r\sin\varphi,\ J = r \\ \Delta = \{(r,\varphi):\ 0 \le \varphi \le \pi \wedge 0 \le r \le \sin\varphi\} \end{array} \right\} = \iint_{\Delta} (6r\cos\varphi - 2) r\,dr\,d\varphi =$$

$$= \int_0^{\pi} \left[\int_0^{\sin\varphi} (6r^2\cos\varphi - 2r)\,dr \right] d\varphi = \int_0^{\pi} (2\sin^3\varphi\cos\varphi - \sin^2\varphi)\,d\varphi =$$

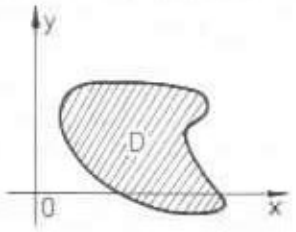
$$= \left[\frac{1}{2}\sin^4\varphi - \frac{1}{2}\varphi + \frac{1}{4}\sin 2\varphi \right]_0^{\pi} = -\frac{\pi}{2}. \quad \blacksquare$$

NIEZALEŻNOŚĆ CAŁKI KRZYWOLINIOWEJ SKIEROWANEJ OD DROGI CAŁKOWANIA.

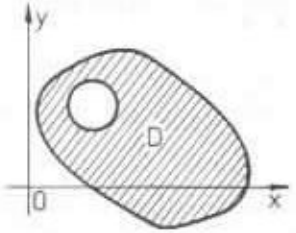
Zdefiniujemy najpierw obszar jednospójny. Obszar D nazywamy **jednospójnym**, gdy wnętrze każdej krzywej zamkniętej zawartej w obszarze D również zawiera się w tym obszarze.

Obszar ograniczony D jest jednospójny, gdy jego dopełnienie jest zbiorem spójnym.

Dla przykładu: obszar na rysunku 7.12 jest jednospójny, natomiast na rysunku 7.13 nie jest jednospójny.



Rys 7.12.



Rys 7.13.

Mówimy, że **całka krzywoliniowa skierowana** pary funkcji P, Q **nie zależy od drogi całkowania** w jednospójnym obszarze D, jeżeli dla każdych dwóch krzywych kawałkami gładkich K_1 i K_2 o wspólnym początku i wspólnym końcu, zawartych w obszarze D zachodzi równość:

$$\int_{K_1} P(x,y)dx + Q(x,y)dy = \int_{K_2} P(x,y)dx + Q(x,y)dy$$

W takim przypadku całkę krzywoliniową skierowaną funkcji P i Q po dowolnej krzywej kawałkami gładkiej zawartej w obszarze D, o początku w punkcie A i końcu w punkcie B, oznaczamy także symbolem

$$\int_A^B P(x,y)dx + Q(x,y)dy .$$

TWIERDZENIE 7.3 Załóżmy, że funkcje P(x,y), Q(x,y) oraz ich pochodne P'_y i Q'_x są ciągłe na jednospójnym obszarze D. Wówczas równoważne są następujące trzy warunki:

(1) $P'_y(x,y) = Q'_x(x,y)$ dla $(x,y) \in D$,

(2) wyrażenie $P(x,y)dx + Q(x,y)dy$ jest różniczką zupełną pewnej funkcji na obszarze D,

(3) całka krzywoliniowa skierowana pary funkcji P, Q nie zależy od drogi całkowania w obszarze D.

Funkcję, dla której wyrażenie $P(x,y)dx + Q(x,y)dy$ jest różniczką zupełną, nazywamy funkcją pierwotną tej różniczki lub funkcją pierwotną pary funkcji P, Q.

Równoważność warunków (1), (2) i (3) oznacza, że każdy z nich jest warunkiem koniecznym i wystarczającym dla każdego z pozostałych.

Udowodnimy implikację $(2) \Rightarrow (3)$:

Niech A i B będą dowolnymi punktami obszaru D i niech K oznacza dowolną krzywą kawałkami gładką zawartą w obszarze D, o początku A i końcu B. Załóżmy, że krzywa K jest określona równaniami

$$x = x(t),\ y = y(t),\ t \in \langle \alpha, \beta \rangle .$$

Zatem $A = (x(\alpha), y(\alpha))$, $B = (x(\beta), y(\beta))$.

Niech F oznacza funkcję, dla której wyrażenie $P(x,y)dx + Q(x,y)dy$ jest różniczką zupełną na obszarze D. Stąd

$$P(x,y) = F'_x(x,y),\quad Q(x,y) = F'_y(x,y) \quad \text{dla } (x,y) \in D .$$

Zatem

$$I \equiv \int_K P(x,y)dx + Q(x,y)dy = \int_K F'_x(x,y)dx + F'_y(x,y)dy =$$

$$= \int_{\alpha}^{\beta} (F'_x(x(t),y(t))x'(t) + F'_y(x(t),y(t))y'(t))dt .$$

Przyjmując oznaczenie $\Phi(t)=F(x(t),y(t))$ mamy

$$\Phi'(t)=F'_x(x(t),y(t))x'(t)+F'_y(x(t),y(t))y'(t).$$

W konsekwencji

$$I=\int_\alpha^\beta \Phi'(t)dt=\Phi(t)\Big|_\alpha^\beta=F(x(t),y(t))\Big|_\alpha^\beta=$$

$$=F(x(\beta),y(\beta))-F(x(\alpha),y(\alpha))=F(B)-F(A),$$

czyli

$$\int_K P(x,y)dx+Q(x,y)dy=F(B)-F(A).$$

Wynik ten oznacza, że całka krzywoliniowa funkcji P i Q nie zależy od drogi całkowania w obszarze D.

Z twierdzenia 7.3, a także z przeprowadzonego tu dowodu, wynika następujący wniosek.

WNIOSEK. Jeżeli funkcje P(x,y), Q(x,y) i ich pochodne P'_y, Q'_x są ciągłe na obszarze jednospójnym D oraz

$$P'_y(x,y)=Q'_x(x,y) \quad \text{dla} \quad (x,y)\in D,$$

to całka krzywoliniowa skierowana funkcji P, Q nie zależy od drogi całkowania w obszarze D, przy czym dla dowolnych punktów A i B tego obszaru mamy

$$\int_A^B P(x,y)dx+Q(x,y)dy=F(B)-F(A)\equiv F(x,y)\Big|_A^B, \tag{7.8}$$

gdzie F oznacza dowolną z funkcji pierwotnych wyrażenia podcałkowego.

PRZYKŁAD 7.4 Obliczymy całkę

$$\int_{(0,1)}^{(3,0)} \ln(1+y^2)dx+\left(\frac{2xy}{1+y^2}-2y\right)dy.$$

Ponieważ

$$P'_y=\frac{2y}{1+y^2}, \quad Q'_x=\frac{2y}{1+y^2}, \quad P'_y=Q'_x \quad \text{dla} \quad (x,y)\in R^2,$$

więc całka ta (tak, jak to wynika z zapisu) nie zależy od drogi całkowania w obszarze $D=R^2$. Całkę tę możemy obliczyć korzystając ze wzoru (7.8) lub, po wybraniu dowolnej drogi całkowania o początku $A=(0,1)$ i końcu $B=(3,0)$, zamieniając ją na całkę oznaczoną.

Metoda 1. Szukamy funkcji pierwotnej wyrażenia podcałkowego, tzn. takiej funkcji F, dla której

$$(1) \quad \begin{cases} F'_x=\ln(1+y^2), \\ F'_y=\dfrac{2xy}{1+y^2}-2y. \end{cases}$$

Z pierwszego z tych równań, przez całkowanie względem x, otrzymujemy, że

$$(2) \quad F(x,y)=x\ln(1+y^2)+\varphi(y),$$

gdzie $\varphi(y)$ jest pewną funkcją zmiennej y. Z równości (2) wynika, że

$$F'_y=\frac{2xy}{1+y^2}+\varphi'(y).$$

Porównując ten wynik z drugim równaniem układu (1) mamy

$$\frac{2xy}{1+y^2}+\varphi'(y)=\frac{2xy}{1+y^2}-2y,$$

$$\varphi'(y)=-2y,$$

$$\varphi(y)=-y^2+C,$$

gdzie C jest dowolną stałą. Tak więc, zgodnie z (2), każda funkcja postaci

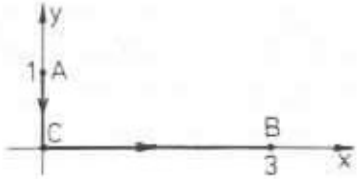
$$F(x,y)=x\ln(1+y^2)-y^2+C,$$

jest funkcją pierwotną wyrażenia podcałkowego.

Korzystając ze wzoru (7.8) mamy

$$\int_{(0,1)}^{(3,0)} \ln(1+y^2)dx+\left(\frac{2xy}{1+y^2}-2y\right)dy=\left(x\ln(1+y^2)-y^2\right)\Big|_{(0,1)}^{(3,0)}=1.$$

Metoda 2. Całkę obliczymy przyjmując, że krzywą całkowania jest łamana ACB, gdzie $C=(0,0)$, (rys. 7.14). Wówczas



Rys 7.14

$$\int_{(0,1)}^{(3,0)} \ln(1+y^2)dx+(\frac{2xy}{1+y^2}-2y)dy = \int_{AC}+\int_{CB} = \begin{Bmatrix} AC:\ x=0,\ y=1-t,\ t\in<0,1> \\ CB:\ x=3t,\ y=0,\ t\in<0,1> \end{Bmatrix} =$$

$$= \int_0^1 (2-2t)dt + \int_0^1 3\ln 1 dt = (2t-t^2)\Big|_0^1 = 1.$$

Przy wybranej przez nas drodze całkowania, obliczenie całki okazało się bardzo proste.

A teraz przyjmijmy, że krzywą całkowania jest odcinek o początku $A=(0,1)$ i końcu $B=(3,0)$. Wówczas

$$\int_{(0,1)}^{(3,0)} \ln(1+y^2)dx+(\frac{2xy}{1+y^2}-2y)dy = \begin{Bmatrix} AB:\ x=3t,\ y=1-t,\ t\in<0,1> \\ x'(t)=3,\ y'(t)=-1 \end{Bmatrix} =$$

$$= \int_0^1 [3\ln(2-2t+t^2) - \frac{6t(1-t)}{2-2t+t^2} + 2(1-t)]dt.$$

Obliczenie otrzymanej całki oznaczonej wymaga, niestety, żmudnych rachunków. Okazuje się więc, że wybór krzywej całkowania ma w tym przypadku ważne znaczenie – może ułatwić lub utrudnić obliczenie całki.

Obliczając całkę krzywoliniową niezależną od drogi całkowania metodą 2 warto pamiętać, że na ogół najlepiej jest przyjąć za krzywą całkowania łamaną o odcinkach równoległych do osi układu współrzędnych. ■

ZADANIA DO ROZWIĄZANIA.

1. Obliczyć całkę krzywoliniową $\int_K xydx+xdy$, jeśli

   a) K: $x=t$, $y=t^2-1$, $t\in<-1,2>$,

   b) K jest łukiem paraboli $y=x^2+x$ od punktu (0,0) do punktu (-2,2),

   c) K: $x=\cos t$, $y=\sin t$, $t\in<0,\pi>$,

   d) K jest łukiem krzywej $y=\ln x$ dla $x\in<1,e>$.

2. Obliczyć całki krzywoliniowe:

   a) $\int_K (2x+3y)dx+(y-1)dy$, jeśli K: $x=t$, $y=t^2$, $t\in<0,2>$,

   b) $\int_K y\ln(x-y)dx-3y^2dy$, jeśli K jest odcinkiem o początku (1,0) i końcu (-1,-2),

   c) $\int_K \frac{x+1}{y}dx-\frac{y-2}{x}dy$, jeśli K jest łamaną ABC, A(1,1), B(2,2), C(1,2),

   d) $\int_K \frac{dx}{1+y}+\frac{dy}{2+x}$, jeśli K jest łamaną ABC, A(1,1), B(3,1), C(2,0),

   e) $\int_K \frac{y^2}{x}dx-\frac{1}{y^2}dy$, jeśli K: $y=\ln x$, $x\in<e,e^2>$,

   f) $\int_K ydx-\frac{x}{\sqrt{x^2+y^2}}dy$, jeśli K jest skierowanym dodatnio okręgiem o równaniu $x^2+y^2=1$,

   g) $\int_K xydy$, jeśli K: $y=\sin x$, $x\in<0,\pi/2>$,

   h) $\int_K (3x^2-y)dx+xdy$, jeśli K jest łukiem elipsy $\frac{x^2}{4}+y^2=1$ zawartym w półpłaszczyźnie $y\geq 0$, o początku (-2,0) i końcu (2,0).

i) $\int_K (x^2+y^2)dx$, jeśli K: $x=\cos t+t\sin t$, $y=\sin t-t\cos t$, $t\in<0,\pi>$

3. Obliczyć całki krzywoliniowe:

a) $\int_K xdx+(x+y)dy+xzdz$, jeśli K jest odcinkiem o początku (-3,0,4) i końcu (0,5,1),

b) $\int_K ydy+xzdz$, jeśli K: $x=\sin t$, $y=\sin t$, $z=t$, $t\in<0,\pi>$,

c) $\int_K x^2dx+(z+x)dy+y^2dz$, jeśli

K: $x=\cos t$, $y=\sin t$, $z=t$, $t\in<0,\pi/2>$.

4. Obliczyć pracę siły $\vec{F}$ na łuku K, jeśli

a) $\vec{F}=x\vec{i}+(x+y)\vec{j}$, jeśli K: $y=x^2$, $0\le x\le 2$,

b) $\vec{F}=[xy,x]$, jeśli K: $y=x^3$, $-1\le x\le 1$,

c) $\vec{F}=[x,-3,x+z]$, jeśli K: $x=t$, $y=t^2$, $z=t^2$, $-1\le t\le 0$.

5. Obliczyć całki krzywoliniowe:

a) $\int_K x^2ydx-y^3dy$, jeśli K jest łamaną ABCA, A(-1,0), B(0,-1), C(0,1),

b) $\int_K (xy^2+e^{-x})dx+(x^2y-e^{-x})dy$, jeśli K jest dodatnio skierowanym brzegiem obszaru $D=\{(x,y): x^2-x\le y\le 5x-x^2\}$,

c) $\int_K (x^2+y)dx+4xydy$, jeśli K jest skierowanym dodatnio okręgiem o równaniu $x^2+y^2=2$,

d) $\int_K \frac{x^2}{2+x}dx+\frac{x}{1+y}dy$, jeśli K jest łamaną ABCDA, A(0,0), B(0,1), C(1,1), D(2,0),

e) $\int_K (x^2-\sqrt{1+x^2})dx+(x^2y+\sqrt{1+y^2})dy$, jeśli K jest skierowanym dodatnio okręgiem o równaniu $x^2+y^2-y=0$,

f) $\int_K (xe^{y^2}-2y)dx+(x^2ye^{y^2}+x)dy$, jeśli K jest skierowanym dodatnio okręgiem o równaniu $x^2+y^2-6x+2y+6=0$,

g) $\int_K (y-\ln(1+x^2))dx+(xy+\ln(1+y^2))dy$ jeśli K jest dodatnio skierowanym brzegiem obszaru $D=\{(x,y): 1\le x^2+y^2\le 4 \wedge y\ge 0\}$,

h) $\int_K (x^2+y^2)dy$, jeśli K jest skierowanym dodatnio brzegiem obszaru ograniczonego krzywymi $y=x^2$, $y=2x-x^2$,

i) $\int_K (x-\sin y)dx+(y+\cos y)dy$, jeśli K jest ujemnie skierowanym brzegiem trójkąta o wierzchołkach (0,0), (1,0), (2,2).

6. Obliczyć całki krzywoliniowe:

a) $\int_{(-1,3)}^{(3,-5)} (2x+y)dx+(x+2y)dy$,

b) $\int_{(2,2)}^{(0,1)} (1+xy)e^{xy}dx+x^2e^{xy}dy$,

c) $\int_K \frac{y^2}{x}dx+(1+2y\ln x)dy$, jeśli K jest dowolną krzywą kawałkami gładką nie przecinającą osi 0y, o początku (1,0) i końcu (2,3),

d) $\int_K \frac{x}{x^2+y^2}dx+\frac{y}{x^2+y^2}dy$, jeśli K jest dowolną krzywą kawałkami gładką zawartą w półpłaszczyźnie $y>0$ o początku (-1,1) i końcu (2,2),

e) $\int_{(0,0)}^{(1,\pi)} 2x\cos y\,dx+(\sin 2y-x^2\sin y)dy$,

f) $\int_K \frac{2y}{(y-x)^2}dx-\frac{2x}{(y-x)^2}dy$, jeśli K jest dowolną krzywą kawałkami gładką nie przecinającą prostej $y=x$, o początku (3,2) i końcu (1,0).

7. Obliczyć całki krzywoliniowe:

a) $\int_K \frac{dx}{2-y^2}+dy$, jeśli K: $x=2\cos t$, $y=\sin t$, $t\in<0,\pi>$,

b) $\int_K (2x-2xye^{-x^2})dx+e^{-x^2}dy$, jeśli K jest dowolną krzywą kawałkami gładką o początku (0,2) i końcu (3,0),

c) $\int_K (x^2-y^2)dx+(x+y^2)dy$, K jest skierowanym dodatnio brzegiem obszaru $D=\{(x,y): x^2+y^2\le 4 \wedge x\ge 0\}$,

d) $\int_{(0,0)}^{(3,-1)} 2\operatorname{arctg} y\,dx+\frac{2x}{1+y^2}dy$,

e) $\int_K (2y+xy^2)dx+x^2ydy$, jeśli K jest dodatnio skierowanym okręgiem o równaniu $x^2+y^2-2x=0$,

f) $\int_{(0,-1)}^{(1,1)} \frac{2xy}{1+x^2}dx+(2y+\ln(1+x^2))dy$,

g) $\int_K 2ydx+(1-y)\ln(1+x^2)dy$, jeśli K jest łamaną ABC, A(-1,0), B(0,1), C(1,0),

h) $\int_K \frac{xy^2-y}{1+x^2}dx+y\ln(x^2+1)dy$, jeśli K jest skierowanym dodatnio brzegiem obszaru ograniczonego parabolą $y=x^2+x$ i prostą $y=x+1$,

i) $\int_K \frac{1}{\sqrt{x^2+y^2}}dy$, jeśli K jest skierowanym dodatnio okręgiem o równaniu $x^2+y^2=4$,

j) $\int_K (3x^2y+e^x)dx+(x^3-2xe^y)dy$, jeśli K jest łamaną ABCA, A(-1,0), B(0,2), C(1,2),

k) $\int_K \frac{dx}{\sqrt{y^2-x^2}}-\frac{xdy}{y\sqrt{y^2-x^2}}$, jeśli K jest dowolną krzywą kawałkami gładką o początku (0,1), końcu (-1,2) i nie przecinającą prostych $y=x$ i $y=-x$,

l) $\int_K \frac{2y-3x}{\sqrt{y-x}}dx+(x+\frac{x}{\sqrt{y-x}})dy$, jeśli K jest skierowanym dodatnio brzegiem obszaru ograniczonego krzywymi $y=1/x^2, y=4, y=9$.

ł) $\int_K dx+x\ln(x^2+y^2-1)dy$, jeśli K jest skierowanym dodatnio okręgiem o równaniu $x^2+y^2=4$.

Odpowiedzi.

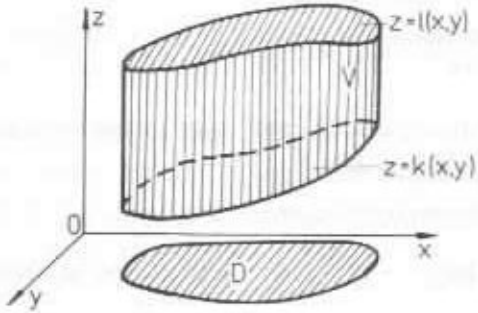
1. a) $-\frac{15}{4}$, b) -2, c) $\frac{\pi}{2}$, d) $\frac{1}{4}(e^2+4e-3)$.
2. a) 16, b) 8, c) $\ln 8-5/4$, d) $1+\ln 0{,}4$, e) $11/6$, f) -2π, g) $\pi/8$, h) $16-2\pi$, i) $10-3\pi^2$.
3. a) 14, b) π, c) $\pi-\frac{4}{3}$.
4. a) 46/3, b) 2/5; c) 8/3.
5. a) $-\frac{1}{6}$, b) $2+10e^{-3}$; c) -2π, d) $1-\ln 8$, e) 0, f) 12π, g) $\frac{14}{3}-\frac{3}{2}\pi$, h) $\frac{1}{3}$, i) $\frac{1}{2}(\cos 2-1)$.
6. a) 12, b) $-2e^4$, c) $3+9\ln 2$, d) $\ln 2$, e) -1, f) 4.
7. a) $-\pi$, b) 7, c) 2π, d) $-\frac{3}{2}\pi$, e) -2π, f) $\ln 2$, g) 2, h) $\pi-2$, i) 0, j) $1+e^2$, k) $-\frac{\pi}{6}$, l) 4, ł) $4\pi\ln 3$.

8. CAŁKA POTRÓJNA.

OKREŚLENIE CAŁKI POTRÓJNEJ. Obszar domknięty $V \subset R^3$ nazywamy **normalnym względem płaszczyzny 0xy**, jeśli

$$V = \{(x,y,z) \in R^3 : k(x,y) \le z \le l(x,y) \wedge (x,y) \in D\},$$

gdzie k i l są funkcjami ciągłymi na regularnym obszarze D oraz $k(x,y) < l(x,y)$ dla punktów wewnętrznych obszaru D (rys. 8.1).



Rys 8.1.

Analogicznie definiuje się obszar V normalny względem płaszczyzny 0yz lub 0xz.

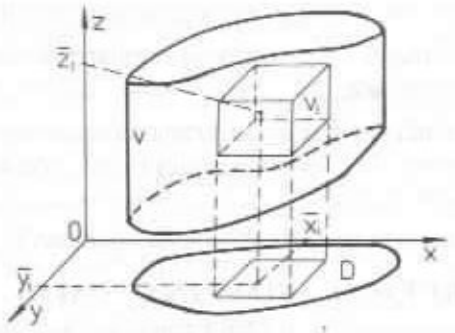
Domknięty obszar $V \subset R^3$ nazywamy **regularnym**, gdy jest on sumą skończonej liczby obszarów normalnych względem jednej z płaszczyzn układu współrzędnych i nie mających wspólnych punktów wewnętrznych.

Niech f będzie funkcją określoną na regularnym obszarze $V \subset R^3$. Podzielmy obszar V w dowolny sposób (np. płaszczyznami prostopadłymi do osi 0x, 0y i 0z) na n obszarów częściowych V_i odpowiednio o objętościach ΔV_i, $i = 1,2,\ldots,n$, tak, by

1° żadne dwa obszary V_i, V_j, $i \neq j$, nie miały wspólnych punktów wewnętrznych,

2° suma obszarów częściowych V_i była obszarem V

$V = V_1 \cup V_2 \cup \cdots \cup V_n$, (rys 8.2).



Rys 8.2.

Analogicznie, jak w definicji całki podwójnej, definiujemy średnicę δ_n podziału oraz sumę

$$S_n = \sum_{i=1}^{n} f(\bar{x}_i, \bar{y}_i, \bar{z}_i)\Delta V_i,$$

gdzie $(\bar{x}_i, \bar{y}_i, \bar{z}_i)$ jest dowolnie wybranym punktem pośrednim w obszarze częściowym V_i, $i = 1,2,\ldots,n$.

Utwórzmy normalny ciąg podziałów obszaru V na obszary częściowe (tzn., że $\delta_n \to 0$ przy $n \to \infty$) i odpowiadający mu ciąg (S_n).

Jeżeli dla każdego normalnego ciągu podziałów obszaru V i każdego wyboru punktów pośrednich $(\bar{x}_i, \bar{y}_i, \bar{z}_i)$ w obszarach częściowych istnieje ta sama skończona granica ciągu (S_n), to tę granicę nazywamy **całką potrójną** funkcji f na obszarze V i oznaczamy symbolem

$$\iiint_V f(x,y,z)\,dxdydz.$$

Zatem

$$\iiint_V f(x,y,z)\,dxdydz \overset{\text{def}}{=} \lim_{n\to\infty} \sum_{i=1}^{n} f(\bar{x}_i, \bar{y}_i, \bar{z}_i)\Delta V_i.$$

Funkcję f, dla której istnieje całka potrójna na obszarze V nazywamy funkcją całkowalną na obszarze V.

TWIERDZENIE 8.1 (warunek konieczny całkowalności). Jeżeli f jest funkcją całkowalną na regularnym obszarze $V \subset R^3$, to jest funkcją ograniczoną na tym obszarze.

TWIERDZENIE 8.2 (warunki wystarczające całkowalności).

(1) Jeżeli funkcja f jest ciągła na regularnym obszarze V, to jest całkowalna na tym obszarze.

(2) Jeżeli funkcja f jest ograniczona na regularnym obszarze V, a jej punkty nieciągłości leżą na skończonej liczbie płatów powierzchniowych o równaniach $z=z(x,y)$, $y=y(x,z)$ lub $x=x(y,z)$ zawartych w obszarze V, to funkcja ta jest całkowalna na obszarze V.

INTERPRETACJA FIZYCZNA CAŁKI POTRÓJNEJ. Niech ρ będzie c i ą g ł ą i n i e u j e m n ą funkcją na regularnym obszarze $V\subset R^3$. Niech $\rho(x,y,z)$ oznacza gęstość masy w punkcie (x,y,z) tego obszaru. Podamy definicję masy obszaru przestrzennego V.

Podzielmy obszar V tak, jak w definicji całki potrójnej. Iloczyn

$$\rho(\overline{x}_i,\overline{y}_i,\overline{z}_i)\Delta V_i, \quad i=1,2,\ldots,n$$

jest równy masie jednorodnego obszaru V_i o objętości ΔV_i i stałej gęstości $\rho(\overline{x}_i,\overline{y}_i,\overline{z}_i)$ i jest w przybliżeniu równy masie obszaru V_i o objętości ΔV_i. Zatem sumę

$$\sum_{i=1}^{n}\rho(\overline{x}_i,\overline{y}_i,\overline{z}_i)\Delta V_i$$

możemy przyjąć za wartość przybliżoną masy obszaru V.

Naturalnym jest więc przyjąć następującą definicję masy obszaru $V\subset R^3$, którego gęstość w każdym punkcie (x,y,z) jest równa $\rho(x,y,z)$:

$$m \overset{def}{=} \lim_{n\to\infty}\sum_{i=1}^{n}\rho(\overline{x}_i,\overline{y}_i,\overline{z}_i)\Delta V_i .$$

Ponieważ ρ jest funkcją ciągłą na obszarze V, więc granica ta istnieje i jest równa całce potrójnej $\iiint_V \rho(x,y,z)\,dxdydz$. Stąd wynika następująca interpretacja fizyczna całki potrójnej:

Całka potrójna funkcji ρ c i ą g ł e j i n i e u j e m n e j na regularnym obszarze $V\subset R^3$ jest równa masie tego obszaru V, którego gęstość w każdym punkcie (x,y,z) jest równa $\rho(x,y,z)$:

$$\text{(8.1)} \qquad \iiint_V \rho(x,y,z)\,dxdydz = m .$$

W szczególności, gdy $\rho(x,y,z)=1$ dla $(x,y,z)\in V$, $m=1\cdot|V|$. Zatem

$$\iiint_V dxdydz = |V| .$$

Dla przykładu: a) całka $\iiint_V dxdydz$, jeśli $V=\{(x,y,z): x^2+y^2+z^2\le 9\}$ jest równa objętości kuli o środku $(0,0,0)$ i promieniu 3; zatem

$$\iiint_V dxdydz = \frac{4}{3}\pi(3)^3 = 36\pi ,$$

b) całka $\iiint_V dxdydz$ jeśli $V=\{(x,y,z): 0\le z\le 6-2x-3y,\ x,y\ge 0\}$ jest równa objętości ostrosłupa o wysokości równej 6 i polu podstawy $\frac{1}{2}\cdot 3\cdot 2=3$ (rys 1.7, rozdz. III); zatem

$$\iiint_V dxdydz = \frac{1}{3}\cdot 6\cdot 3 = 6 .$$

OBLICZANIE CAŁKI POTRÓJNEJ. Całki potrójne obliczamy korzystając z następującego twierdzenia:

TWIERDZENIE 8.3. Jeżeli f jest funkcją ciągłą na obszarze V normalnym względem płaszczyzny 0xy:

$$V=\{(x,y,z)\in R^3: k(x,y)\le z\le l(x,y)\ \wedge\ (x,y)\in D\} ,$$

to

$$\text{(8.2)} \qquad \iiint_V f(x,y,z)\,dxdydz = \iint_D \Big[\int_{k(x,y)}^{l(x,y)} f(x,y,z)\,dz\Big]dxdy .$$

Jeśli ponadto obszar D jest normalny względem osi 0x:

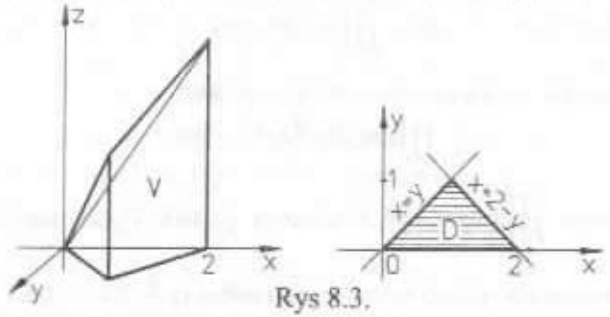
$$D=\{(x,y): a\le x\le b,\ g(x)\le y\le h(x)\} ,$$

to

$$\iiint_V f(x,y,z)\,dxdydz = \int_a^b \Big\{\int_{g(x)}^{h(x)}\Big[\int_{k(x,y)}^{l(x,y)} f(x,y,z)\,dz\Big]dy\Big\}dx .$$

Analogiczne twierdzenia są prawdziwe, gdy obszar całkowania V jest normalny względem innej płaszczyzny układu współrzędnych.

P R Z Y K Ł A D 8.1. Obliczymy całkę $\iiint_V (x-2z)dxdydz$, jeśli $V=\{(x,y,z)\in R^3: 0\le z\le 4x \wedge (x,y)\in D\}$, przy czym D oznacza trójkąt ograniczony prostymi $y=x$, $y=2-x$, $y=0$, (rys 8.3).



Rys 8.3.

Zgodnie z twierdzeniem 8.3 mamy:

$$\iiint_V (x-2z)dxdydz = \iint_D \left[\int_0^{4x}(x-2z)dz\right]dxdy = \iint_D (xz-z^2)\Big|_{z=0}^{z=4x} dxdy =$$

$$= -12\iint_D x^2 dxdy = \left\{D=\{(x,y):\ 0\le y\le 1 \wedge y\le x\le 2-y\}\right\} =$$

$$= -12\int_0^1\left[\int_y^{2-y} x^2 dx\right]dy = -14. \quad \blacksquare$$

P R Z Y K Ł A D 8.2. Obliczymy całkę $\iiint_V 3z\,dxdydz$, jeśli

$$V=\{(x,y,z)\in R^3:\ -1\le x\le\sqrt{y^2+z^2} \wedge (y,z)\in D\},$$

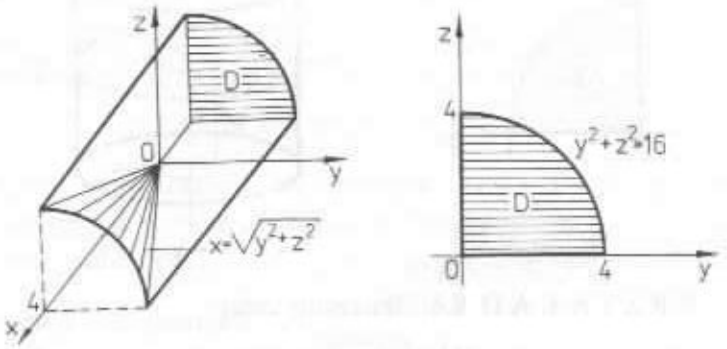
przy czym $D=\{(y,z)\in R^2: y^2+z^2\le 16 \wedge y\ge 0 \wedge z\ge 0\}$, (rys 8.4).

Obszar całkowania V jest normalny względem płaszczyzny 0yz. Skorzystamy z twierdzenia analogicznego do twierdzenia 8.3. Zatem

$$\iiint_V 3z\,dxdydz = \iint_D \left[\int_{-1}^{\sqrt{y^2+z^2}} 3z\,dx\right]dydz = 3\iint_D z(\sqrt{y^2+z^2}+1)dydz =$$

$$= \left\{\begin{matrix} y=r\cos\varphi,\ z=r\sin\varphi,\ J=r \\ \Delta=\{(r,\varphi):\ 0\le r\le 4 \wedge 0\le\varphi\le\pi/2\}\end{matrix}\right\} = 3\iint_\Delta r\sin\varphi\,(r+1)r\,drd\varphi =$$

$$= 3\int_0^4\left[\int_0^{\pi/2}(r^3+r^2)\sin\varphi\,d\varphi\right]dr = 3\int_0^4 (r^3+r^2)dr = 256. \quad \blacksquare$$



Rys 8.4.

P R Z Y K Ł A D 8.3. Obliczymy masę m bryły określonej nierównościami: $x^2+y^2+z^2\le 1$, $x\ge 0$, $y\ge 0$, jeśli gęstość w dowolnym punkcie (x,y,z) jest równa $\rho(x,y,z)=x^2+y^2$.

Zgodnie z (8.1) mamy

$$m = \iiint_V (x^2+y^2)dxdydz.$$

gdzie

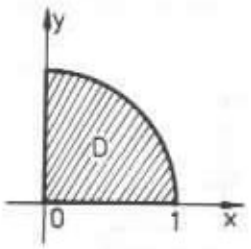
$$V=\{(x,y,z)\in R^3: -\sqrt{1-x^2-y^2}\le z\le\sqrt{1-x^2-y^2} \wedge (x,y)\in D\},$$

przy czym $D=\{(x,y)\in R^2: x^2+y^2\le 1 \wedge x\ge 0 \wedge y\ge 0\}$ (rys. 8.5). Zatem

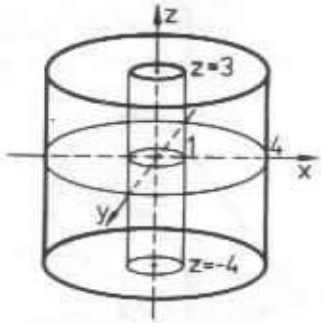
$$m = \iint_D\left[\int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}}(x^2+y^2)dz\right]dxdy = \iint_D 2(x^2+y^2)\sqrt{1-x^2-y^2}\,dxdy =$$

$$= \left\{\begin{matrix} x=r\cos\varphi,\ y=r\sin\varphi,\ J=r \\ \Delta=\{(r,\varphi):\ 0\le r\le 1 \wedge 0\le\varphi\le\pi/2\}\end{matrix}\right\} = \iint_\Delta 2r^2\sqrt{1-r^2}\,r\,drd\varphi =$$

$$= \int_0^1 [\int_0^{\pi/2} 2r^3\sqrt{1-r^2}\,d\varphi]\,dr = \frac{\pi}{2}\int_0^1 2r^3\sqrt{1-r^2}\,dr = \frac{2}{15}\pi . \quad \blacksquare$$



Rys 8.5.



Rys 8.6.

PRZYKŁAD 8.4. Obliczymy całkę

$$\iiint_V z\sqrt{x^2+y^2}\,dxdydz ,$$

jeśli V jest obszarem przestrzennym ograniczonym powierzchniami: $z=-4$, $z=3$, $x^2+y^2=1$, $x^2+y^2=16$, (rys 8.6).

Ponieważ

$$V=\{(x,y,z)\in R^3:\ -4\le z\le 3 \wedge (x,y)\in D\},$$

$$D=\{(x,y)\in R^2: 1\le x^2+y^2\le 16\},$$

więc

$$\iiint_V z\sqrt{x^2+y^2}\,dxdydz = \iint_D [\int_{-4}^{3} z\sqrt{x^2+y^2}\,dz]dxdy = -\frac{7}{2}\iint_D \sqrt{x^2+y^2}\,dxdy =$$

$$= \left\{\begin{matrix} x=r\cos\varphi,\ y=r\sin\varphi,\ J=r \\ \Delta=\{(r,\varphi): 1\le r\le 4 \wedge 0\le\varphi\le 2\pi\}\end{matrix}\right\} = -\frac{7}{2}\iint_\Delta r^2\,dr\,d\varphi = -\frac{7}{2}\int_1^4 [\int_0^{2\pi} r^2\,d\varphi]\,dr =$$

$$= -7\pi\int_1^4 r^2\,dr = -147\pi . \quad \blacksquare$$

ZADANIA DO ROZWIĄZANIA.

1. Korzystając z całki potrójnej obliczyć objętość bryły V ograniczonej powierzchniami:
   a) $z=-2$, $z=\sqrt{x^2+y^2}$, $x^2+y^2=9$,
   b) $z=0$, $z=x+y-1$, $x^2+y^2+2y=0$.
2. Korzystając z całki potrójnej obliczyć objętość bryły V, jeśli:
   a) $V=\{(x,y,z)\in R^3: x\ge 0, y\ge 0 \wedge x^2+y^2\le 9 \wedge 0\le z\le x+y+1\}$,
   b) $V=\{(x,y,z)\in R^3: x\le y \wedge 0\le z\le 4-x^2-y^2\}$.
3. Obliczyć masę półkuli $V=\{(x,y,z)\in R^3: z\ge 0 \wedge x^2+y^2+z^2\le 16\}$, której gęstość w każdym punkcie (x,y,z) jest równa $\rho(x,y,z)$, jeśli:
   a) $\rho(x,y,z)=z$, b) $\rho(x,y,z)=z^2$, c) $\rho(x,y,z)=y^2$.
4. Obliczyć całki potrójne:
   a) $\iiint_V xydxdydz$, gdy $V=\{(x,y,z)\in R^3: x,y\ge 0 \wedge 0\le z\le 6-2x-y\}$,
   b) $\iiint_V zdxdydz$, gdy

   $V=\{(x,y,z)\in R^3: 0<z\le\sqrt{x^2+y^2} \wedge x^2+y^2+2x\le 0\}$,
   c) $\iiint_V (x+y)dxdydz$, jeśli $V=\{(x,y,z)\in R^3: x^2+y^2-11\le z\le -2\}$,
   d) $\iiint_V (x+y)^2\,dxdydz$, jeśli $V=\{(x,y,z)\in R^3: -2\le z\le 7-x^2-y^2\}$,
   e) $\iiint_V zdxdydz$, jeśli $V=\{(x,y,z)\in R^3: y^2\le x\le 6-y \wedge 0\le z\le\sqrt{x}\}$.

Odpowiedzi.

1. a) 36π, b) 2π, wsk. $V=\{(x,y,z): x+y-1\le z\le 0 \wedge x^2+y^2+2y\le 0\}$
2. a) $18+\frac{9}{4}\pi$, b) 4π.
3. a) 64π, b) $\frac{2\pi}{15}\cdot 4^5$, c) $\frac{2\pi}{15}\cdot 4^5$.
4. a) $\frac{81}{5}$, b) $\frac{3}{4}$, c) 0, d) $\frac{243}{2}\pi$, e) $\frac{125}{3}$.

IV. RÓWNANIA RÓŻNICZKOWE ZWYCZAJNE

1. PODSTAWOWE WIADOMOŚCI.

RÓWNANIA RÓŻNICZKOWE RZĘDU PIERWSZEGO.

Równanie postaci

(1.1) $$F(x,y,y')=0,$$

gdzie F jest funkcją ciągłą na pewnym obszarze $\Delta \subset R^3$, y oznacza **niewiadomą funkcję** zmiennej x określoną na pewnym przedziale, a y' - pochodną tej niewiadomej funkcji, nazywamy **równaniem różniczkowym zwyczajnym rzędu pierwszego** pod warunkiem, że y' występuje w tym równaniu w sposób wyraźny.

Na przykład równania

$$y'-2y^2=0, \qquad xy'-y(\ln y-\ln x)=0, \qquad (y')^2-4x^2=0$$

są równaniami różniczkowymi zwyczajnymi I rzędu.

Jeżeli równanie (1.1) można rozwiązać ze względu na pochodną y', wówczas równanie to można zapisać w tzw. **postaci normalnej**

(1.2) $$y'=f(x,y),$$

gdzie f jest funkcją ciągłą na pewnym obszarze $D \subset R^2$.

W dalszym ciągu będziemy rozważać równania różniczkowe zapisane na ogół w postaci normalnej, a za obszar D będziemy przyjmować, o ile nie będzie innej informacji, dziedzinę naturalną funkcji f.

Funkcja niewiadoma w równaniach (1.1) i (1.2) jest funkcją jednej zmiennej, zatem pochodna, która w nich występuje jest pochodną zwyczajną funkcji. Dlatego właśnie równania te nazywamy równaniami różniczkowymi zwyczajnymi, w odróżnieniu od równań różniczkowych cząstkowych, w których występują pochodne cząstkowe niewiadomej funkcji wielu zmiennych. Ponieważ w dalszym ciągu będziemy rozważać jedynie równania różniczkowe zwyczajne, będziemy o nich mówić krótko: równania różniczkowe.

KLASYFIKACJA ROZWIĄZAŃ. Każdą funkcję $y=y(x)$ różniczkowalną na przedziale I, o wykresie zawartym w obszarze D (tzn. $(x,y(x)) \in D$ dla $x \in I$), taką, że

$$\bigwedge_{x\in I} \; y'(x)=f(x,y(x))$$

nazywamy **rozwiązaniem (całką)** równania (1.2) na przedziale I.

Rozwiązanie tu zdefiniowane nazywamy też **rozwiązaniem szczególnym (całką szczególną)** równania różniczkowego, w odróżnieniu od rozwiązania ogólnego (całki ogólnej), które zdefiniujemy później.

Wykres rozwiązania $y=y(x)$, $x \in I$, równania różniczkowego nazywamy **krzywą całkową** tego równania. Znajdowanie rozwiązań równania różniczkowego nazywamy rozwiązywaniem lub całkowaniem tego równania.

PRZYKŁAD 1.1. Sprawdzimy, że funkcja

(1) $$y(x)=x+x^2-e^{2x}, \quad x \in R,$$

jest rozwiązaniem równania różniczkowego

(2) $$y'=2y-2x^2+1$$

na obszarze $D=R^2$.

Funkcja (1) jest funkcją różniczkowalną w każdym punkcie $x \in R$, a jej wykres zawiera się oczywiście w obszarze $D=R^2$. Obliczamy pochodną $y'(x)$ funkcji (1), a następnie $y(x)$ i $y'(x)$ podstawiamy do równania (2). Ponieważ lewa strona tego równania

$$L=y'(x)=1+2x-2e^{2x},$$

a prawa strona

$$P = 2y(x) - 2x^2 + 1 = 2x + 2x^2 - 2e^{2x} - 2x^2 + 1 = 1 + 2x - 2e^{2x},$$

więc dla funkcji $y = y(x)$ określonej wzorem (1) równanie (2) jest spełnione dla każdego $x \in R$. Funkcja (1) jest rozwiązaniem równania różniczkowego (2). ■

PRZYKŁAD 1.2. Weźmy pod uwagę równanie

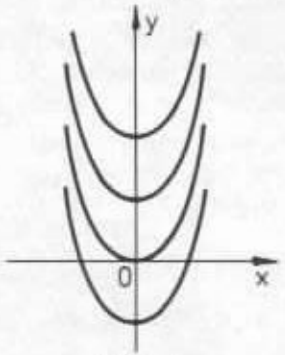
(1) $$y' = 2x.$$

Równanie rozważamy w dziedzinie naturalnej funkcji $f(x,y) = 2x$, tzn. dla $(x,y) \in D = R^2$. Każda funkcja postaci

(2) $$y = x^2 + C, \quad x \in I = (-\infty, +\infty), \ C \in R,$$

jest rozwiązaniem równania (1), przy czym innych rozwiązań określonych na przedziale I równanie nie ma.

Rozwiązania te stanowią jednoparametrową rodzinę funkcji, a ich wykresy (krzywe całkowe równania (1)) są jednoparametrową rodziną parabol, z których kilka przedstawiono na rysunku 1.1.



Rys 1.1.

Mówimy, że wzór (2) określa rozwiązanie ogólne (całkę ogólną) równania (1).

Zauważmy, że przez każdy punkt płaszczyzny przechodzi dokładnie jedna krzywa całkowa rozważanego równania. Wyznaczymy na przykład tę krzywą całkową, która przechodzi przez punkt $P_0(1,2)$. Oznacza to, że należy tak dobrać C we wzorze (2), by dla $x = 1$ otrzymać $y = 2$. Podstawiając $x = 1$ i $y = 2$ we wzorze (2) otrzymujemy $C = 1$.

Krzywa całkowa spełniająca podany warunek jest więc określona wzorem

$$y = x^2 + 1, \quad x \in I = (-\infty, +\infty).$$

Inaczej mówiąc, znaleźliśmy to rozwiązanie równania (1) na przedziale I, które spełnia tzw. warunek początkowy $y(1) = 2$.

Dodajmy, że rozwiązaniem równania (1) jest także każda funkcja określona wzorem

$$y = x^2 + C, \quad x \in (\alpha, \beta), \ C \in R,$$

gdzie (α,β) jest dowolnym przedziałem zawartym w przedziale $I = (-\infty, +\infty)$. ■

Załóżmy, że funkcje

(a) $$y = y_1(x), \ x \in I_1 \quad \text{oraz} \quad \text{(b)} \quad y = y_2(x), \ x \in I_2$$

są rozwiązaniami równania różniczkowego (1.2), przy czym $I_1 \subset I_2$ oraz $y_1(x) = y_2(x)$ dla każdego $x \in I_1$.

Rozwiązanie (a) nazywamy **zawężeniem** rozwiązania (b) do przedziału I_1, a rozwiązanie (b) nazywamy **przedłużeniem** rozwiązania (a) na przedział I_2, przy czym (b) nazywamy **przedłużeniem właściwym** rozwiązania (a), gdy $I_1 \subset I_2$ i $I_1 \neq I_2$.

Rozwiązanie

(c) $$y = y(x), \quad x \in I,$$

równania różniczkowego nazywamy **rozwiązaniem globalnym**, gdy nie istnieje rozwiązanie tego równania, które jest przedłużeniem właściwym rozwiązania (c).

Na przykład funkcja

$$y = x^2 - 1, \quad x \in I = (-\infty, +\infty),$$

jest rozwiązaniem globalnym równania (1) rozpatrywanego w przykładzie 1.2. Łatwo widać, że każde zawężenie tego rozwiązania, czyli funkcja

$$y = x^2 - 1, \quad x \in (\alpha, \beta),$$

gdzie (α,β) jest przedziałem zawartym w przedziale $I = (-\infty, +\infty)$ jest także rozwiązaniem tego równania. Wzór (2) w przykładzie 1.2 określa wszystkie rozwiązania globalne równania (1).

Rozwiązując równanie różniczkowe wyznaczamy z reguły rozwiązania globalne. Pamiętajmy jednak, że z definicji rozwiązania wynika, że każde zawężenie takiego rozwiązania jest także rozwiązaniem rozpatrywanego równania.

Rozwiązaniem ogólnym (całką ogólną) równania (1.2) nazywamy jednoparametrową rodzinę funkcji

$$y = y(x,C)$$

takich, że dla każdej dopuszczalnej wartości C_0 funkcja $y = y(x,C_0)$ jest rozwiązaniem równania (1.2) na pewnym przedziale.

Dodajmy, że w pewnych przypadkach otrzymujemy rozwiązania równania różniczkowego w postaci uwikłanej - rozwiązanie szczególne $y = y(x)$ może być określone równaniem $\varphi(x,y) = 0$, a rozwiązanie ogólne - równaniem postaci $\Phi(x,y,C) = 0$.

Warto jeszcze zauważyć, że rozwiązanie ogólne nie musi zawierać wszystkich rozwiązań równania, choć dla pewnych równań tak jest.

Zadanie polegające na wyznaczeniu rozwiązania $y = y(x)$ równania różniczkowego (1.2) spełniającego tzw. warunek początkowy $y(x_0) = y_0$, gdzie $(x_0,y_0) \in D$, nazywamy **zagadnieniem początkowym** lub inaczej **zagadnieniem Cauchy'ego**.

Geometrycznie zadanie to sprowadza się do znalezienia krzywej całkowej równania różniczkowego przechodzącej przez podany punkt płaszczyzny. Oczywiście może się zdarzyć, że przez dany punkt przechodzi więcej niż jedna krzywa całkowa danego równania i wtedy mówimy, że zagadnienie Cauchy'ego nie ma jednoznacznego rozwiązania.

Rozwiązanie $y = y(x)$, $x \in I$, równania różniczkowego nazywamy **rozwiązaniem regularnym** tego równania na przedziale I, gdy przez żaden punkt krzywej całkowej wyznaczonej przez to rozwiązanie nie przechodzi inna krzywa całkowa tego równania.

Rozwiązanie $y = y(x)$, $x \in I$, równania różniczkowego nazywamy **rozwiązaniem osobliwym** tego równania na przedziale I, gdy przez każdy punkt krzywej całkowej wyznaczonej przez to rozwiązanie przechodzi co najmniej jeszcze jedna, inna krzywa całkowa tego równania.

Na przykład rozwiązania równania $y' = 2x$ rozważanego w przykładzie 1.1 są rozwiązaniami regularnymi na przedziale $I = (-\infty,+\infty)$. Nietrudno bowiem sprawdzić, że przez dowolny punkt $(x_0,y_0) \in R^2$ przechodzi dokładnie jedna krzywa całkowa równania, a mianowicie - parabola $y = x^2 + C$, gdzie $C = y_0 - x_0^2$.

W przykładzie 1.3 będziemy rozważać równanie, które ma rozwiązanie osobliwe.

PRZYKŁAD 1.3. Weźmy pod uwagę równanie

(1) $$y' + 3\sqrt[3]{y^2} = 0, \qquad (x,y) \in D = R^2.$$

Łatwo sprawdzić, że funkcja stała

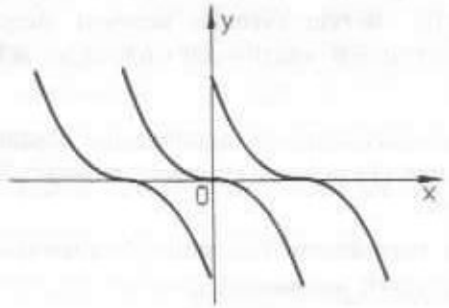
(2) $$y = 0, \qquad x \in I = (-\infty,+\infty),$$

oraz funkcje postaci

(3) $$y = (C - x)^3, \quad x \in I = (-\infty,+\infty),$$

gdzie C jest dowolną stałą, są rozwiązaniami równania (1).

Każdą z krzywych całkowych określonych wzorem (3), a jest ich nieskończenie wiele, można otrzymać przez przesunięcie o wektor [C,0] krzywej $y = -x^3$. Kilka tych krzywych przedstawia rysunek 1.2.



Rys 1.2.

Łatwo stwierdzić, że przez każdy ustalony punkt $(x_0,0)$ krzywej całkowej $y = 0$ przechodzi również krzywa całkowa o równaniu

$y=(C-x)^3$, $C=x_0$, $x\in R$. Zgodnie z przyjętym określeniem rozwiązanie $y=0$, $x\in R$, jest rozwiązaniem osobliwym równania (1).

Natomiast rozwiązanie postaci $y=(C-x)^3$, $x\in R$, dla dowolnej stałej C, nie jest rozwiązaniem regularnym, ani też rozwiązaniem osobliwym.

Zauważmy jeszcze, że wzory (2) i (3) nie określają wszystkich rozwiązań równania (1) na przedziale $I=(-\infty,+\infty)$. Rozwiązaniami tego równania będą również funkcje (a będzie ich nieskończenie wiele) odpowiednio "sklejone" z funkcji (2) i (3), na przykład:

$$y=\begin{cases}0, & x\in(-\infty,1>,\\ (1-x)^3, & x\in(1,+\infty),\end{cases}\qquad y=\begin{cases}(-1-x)^3, & x\in(-\infty,-1>,\\ 0, & x\in(-1,0),\\ -x^3, & x\in<0,+\infty).\end{cases}$$

Czytelnikowi pozostawiamy naszkicowanie krzywych całkowych określonych tymi wzorami.

Jak widać problem zapisania wszystkich rozwiązań równania (1) nawet na ustalonym przedziale $(-\infty,+\infty)$ jest bardziej skomplikowany, niż by się to wydawało na początku.

Można natomiast powiedzieć, że rysunek, na którym znalazłyby się w s z y s t k i e krzywe $y=(C-x)^3$, $x\in R$, $C\in R$, oraz krzywa całkowa $y=0$, $x\in R$, przedstawiałby niewątpliwie wszystkie krzywe całkowe równania (1). Na tym rysunku bowiem znajdowałyby się również wykresy wszystkich możliwych rozwiązań sklejonych z rozwiązań (2) i (3).

Na koniec zauważmy, że ograniczenie obszaru, w którym rozwiązujemy równanie (1) pozwala problem uprościć. Z powyższych rozważań wynika, że:

a) Każde rozwiązanie równania (1) rozważanego na obszarze $D_1=\{(x,y)\in R^2:\ y>0\}$ jest postaci:

$$y=(C-x)^3,\qquad x\in(-\infty,C),\quad C\in R,$$

przy czym przez każdy punkt $(x_0,y_0)\in D_1$ przechodzi dokładnie jedna krzywa całkowa równania (1) i jest ona określona wzorem

$$y=(C_0-x)^3\qquad x\in(-\infty,C_0),\quad C_0=x_0+\sqrt[3]{y_0}.$$

b) Każde rozwiązanie równania (1) rozważanego na obszarze $D_2=\{(x,y)\in R^2:\ y<0\}$ ma postać:

$$y=(C-x)^3,\qquad x\in(C,+\infty),\quad C\in R,$$

przy czym przez każdy punkt $(x_0,y_0)\in D_2$ przechodzi dokładnie jedna krzywa całkowa równania (1) i jest ona określona wzorem

$$y=(C_0-x)^3,\qquad x\in(C_0,+\infty),\quad C_0=x_0+\sqrt[3]{y_0}.$$

c) Ponadto rozwiązaniem równania (1) jest funkcja stała $y=0$, $x\in R$.

Dodajmy, że rozwiązania przebiegające w obszarach D_1 i D_2 są rozwiązaniami regularnymi. Mówimy, że w obszarach D_1 i D_2 dla rozpatrywanego równania ma miejsce jednoznaczność rozwiązań. ■

WARUNKI ISTNIENIA I JEDNOZNACZNOŚCI ROZWIĄZAŃ. Niżej podamy dwa twierdzenia o istnieniu i jednoznaczności rozwiązań dla równań różniczkowych pierwszego rzędu w postaci normalnej:

$$y'=f(x,y).$$

TWIERDZENIE 1.1 (Peano). Jeżeli funkcja f jest ciągła na obszarze $D\subset R^2$, to przez każdy punkt tego obszaru przechodzi p r z y n a j m n i e j j e d n a krzywa całkowa równania $y'=f(x,y)$.

TWIERDZENIE 1.2 (Cauchy'ego). Jeżeli funkcja f jest ciągła i ma ciągłą pochodną f'_y na obszarze D, to przez każdy punkt tego obszaru przechodzi d o k ł a d n i e j e d n a krzywa całkowa równania $y'=f(x,y)$.

P R Z Y K Ł A D 1.4. Weźmy pod uwagę równanie

(1) $$y'=x(y^2-1).$$

Prawa strona tego równania $f(x,y)=x(y^2-1)$ jest funkcją ciągłą na całej płaszczyźnie, zatem z twierdzenia Peano wynika, że przez każdy punkt płaszczyzny przechodzi przynajmniej jedna krzywa całkowa równania (1).

Ponieważ pochodna $f'_y=2xy$ jest także funkcją ciągłą na płaszczyźnie, więc możemy ponadto stwierdzić, że rozwiązania te są regularne

– z twierdzenia Cauchy'ego wynika bowiem, że przez każdy punkt płaszczyzny przechodzi dokładnie jedna krzywa całkowa równania (1). ■

RÓWNANIA RÓŻNICZKOWE WYŻSZYCH RZĘDÓW.

Równanie postaci

$$F(x,y,y',\dots,y^{(n)})=0, \tag{1.3}$$

gdzie F jest funkcją ciągłą na pewnym obszarze $\Delta\subset R^{n+2}$, y oznacza niewiadomą funkcję zmiennej x określoną na pewnym przedziale, a $y',\dots,y^{(n)}$ - kolejne pochodne tej niewiadomej funkcji, przy czym $y^{(n)}$ występuje w tym równaniu w sposób wyraźny, nazywamy **równaniem różniczkowym zwyczajnym rzędu n**.

Tak więc, rząd równania różniczkowego jest to najwyższy z rzędów pochodnych funkcji niewiadomej występujących w tym równaniu. Na przykład równania

$$y''-ye^x=1,\quad y'''+2xy'=0,\quad yy^{(5)}+y'=0$$

są równaniami różniczkowymi odpowiednio drugiego, trzeciego i piątego rzędu.

Jeżeli równanie (1.3) można rozwiązać względem pochodnej najwyższego rzędu, to równanie to można zapisać w tzw. **postaci normalnej**

$$y^{(n)}=f(x,y,y',\dots,y^{(n-1)}), \tag{1.4}$$

gdzie f jest funkcją ciągłą na pewnym obszarze $D\subset R^{n+1}$. Rozważania w tym rozdziale będą dotyczyły przede wszystkim równań w postaci normalnej.

Każdą funkcję $y=y(x)$ n-krotnie różniczkowalną na przedziale I taką, że $(x,y(x),y'(x),\dots,y^{(n-1)}(x))\in D$ dla $x\in I$ oraz

$$\bigwedge_{x\in I} y^{(n)}(x)=f(x,y(x),y'(x),\dots,y^{(n-1)}(x))$$

nazywamy **rozwiązaniem (całką)** równania (1.4).

Rozwiązanie takie nazywamy również **rozwiązaniem szczególnym (całką szczególną)** równania, w odróżnieniu od rozwiązania ogólnego (całki ogólnej), które definiujemy niżej:

Rozwiązaniem ogólnym (całką ogólną) równania (1.4) nazywamy n-parametrową rodzinę funkcji

$$y=y(x,C_1,\dots,C_n)$$

taką, że dla każdego układu stałych $C_1^0,\dots,C_n^0$ z pewnego obszaru zawartego w R^n funkcja

$$y=y(x,C_1^0,\dots,C_n^0)$$

jest rozwiązaniem równania (1.4) na pewnym przedziale.

Zadanie polegające na wyznaczeniu takiego rozwiązania $y=y(x)$ równania (1.4), dla którego spełnione są tzw. warunki początkowe

$$y(x_0)=y_0,\quad y'(x_0)=y_0',\ \dots\ ,y^{(n-1)}(x_0)=y_0^{(n-1)},$$

gdzie $(x_0,y_0,y_0',\dots,y_0^{(n-1)})\in D$, nazywamy **zagadnieniem początkowym** lub inaczej **zagadnieniem Cauchy'ego**.

Dla równania różniczkowego drugiego rzędu zagadnienie Cauchy'ego oznacza wyznaczenie rozwiązania spełniającego warunki

$$y(x_0)=y_0,\quad y'(x_0)=y_0',$$

a geometrycznie - wyznaczenie krzywej całkowej przechodzącej przez dany punkt i mającej w tym punkcie styczną o podanym kierunku.

Zagadnienie brzegowe dla równania różniczkowego rzędu n jest to zadanie polegające na znalezieniu takiego rozwiązania, które w podanych punktach: $x_1,\dots,x_n$ przyjmuje zadane wartości: $y_1,\dots,y_n$, tzn.

$$y(x_1)=y_1,\dots,y(x_n)=y_n.$$

Na przykład dla równania różniczkowego rzędu drugiego zagadnienie brzegowe polega na wyznaczeniu krzywej całkowej, która przechodzi przez dwa podane punkty.

PRZYKŁAD 1.5. Rozwiążemy równanie

$$y'''=6, \tag{1}$$

a następnie znajdziemy rozwiązanie szczególne spełniające

a) warunki brzegowe $y(0)=1,\ y(2)=5,\ y(-2)=-3$,

b) warunki początkowe $y(1)=-1,\ y'(1)=2,\ y''(1)=4$.

Rozwiązanie ogólne wyznaczymy przez trzykrotne całkowanie. Z (1) otrzymujemy kolejno:

$$y'' = 6x + C_1, \quad y' = 3x^2 + C_1x + C_2, \quad y = x^3 + \frac{1}{2}C_1x^2 + C_2x + C_3.$$

Zatem rozwiązanie ogólne danego równania jest trójparametrową rodziną funkcji

$$(2) \qquad y = x^3 + \frac{1}{2}C_1x^2 + C_2x + C_3, \quad x \in R, \quad C_1, C_2, C_3 \in R.$$

a) Uwzględniając warunki brzegowe w (2) otrzymujemy

$$1 = C_3, \quad 5 = 8 + 2C_1 + 2C_2 + C_3, \quad -3 = -8 + 2C_1 - 2C_2 + C_3,$$

a stąd

$$C_1 = 0, \quad C_2 = -2, \quad C_3 = 1.$$

Tak więc, szukane rozwiązanie szczególne jest postaci

$$y = x^3 - 2x + 1, \quad x \in R.$$

b) Dla rozwiązań określonych wzorem (2) mamy

$$y' = 3x^2 + C_1x + C_2, \quad y'' = 6x + C_1.$$

Uwzględniając warunki początkowe otrzymujemy

$$-1 = 1 + \frac{1}{2}C_1 + C_2 + C_3, \quad 2 = 3 + C_1 + C_2, \quad 4 = 6 + C_1,$$

a stąd

$$C_1 = -2, \quad C_2 = 1, \quad C_3 = -2.$$

Zatem rozwiązanie szczególne spełniające podane warunki początkowe określone jest wzorem

$$y = x^3 - x^2 + x - 2, \quad x \in R. \qquad \blacksquare$$

W dalszy ciągu zapoznamy się z pewnymi efektywnymi metodami całkowania niektórych typów równań różniczkowych. Będą to metody, które prowadzą do napisania wzoru na rozwiązanie ogólne (czasem w postaci uwikłanej) danego równania różniczkowego. Niemniej, dla wielu równań takich metod całkowania nie znamy i wtedy musimy poprzestać na znalezieniu przybliżonych rozwiązań tych równań. Zresztą, z punktu widzenia zastosowań praktycznych, na ogół wystarcza w zupełności, gdy możemy podać, z ustaloną dokładnością, wartości przybliżone rozwiązania.

ZADANIA DO ROZWIĄZANIA.

1. Podać po trzy przykłady równań różniczkowych I i II rzędu.
2. Sprawdzić, że funkcja $y = \sin x$, $x \in (0, \pi)$, jest rozwiązaniem równania $y' = 2y\operatorname{ctg} x - \cos x$.
3. Sprawdzić, że funkcja $y = 2e^{x^2} - e^{-x^2}$, $x \in R$, jest rozwiązaniem równania $y' = 2xy + 4xe^{-x^2}$.
4. Sprawdzić, że funkcja $y = x + \frac{1}{\ln x}$, $x > 1$, jest rozwiązaniem równania $xy' + (y - x)^2 - x = 0$ i wykazać, że jest to jedyne rozwiązanie tego równania spełniające warunek $y(e) = 1 + e$.
5. Sprawdzić, że funkcja $y = \frac{1}{x}$, $x > 0$, jest rozwiązaniem równania $y' + \frac{xy^3 + x^2y^4}{1 + y^2} = 0$ i wykazać, że jest to jedyne rozwiązanie tego równania spełniające warunek $y(1) = 1$.
6. Sprawdzić, że funkcja $y = x\sqrt{4 - x^2}$, $x \in (0, 2)$, jest rozwiązaniem równania $y' = \frac{y^2 - x^4}{xy}$ i wykazać, że jest to jedyne rozwiązanie tego równania spełniające warunek $y(1) = \sqrt{3}$.
7. Znaleźć rozwiązanie ogólne i naszkicować kilka krzywych całkowych równania:

   a) $y' = 2x - 1$, b) $y' = 3x^2$, c) $y' = e^x$, d) $y' = 2\cos 2x$,

   e) $y' = \frac{1}{x + 1}$, f) $y' = \frac{1}{x^2}$, g) $y' = \frac{1}{x^3}$, h) $y' = \frac{1}{2\sqrt{x}}$.
8. Znaleźć rozwiązanie równania spełniające podany warunek początkowy i naszkicować odpowiednią krzywą całkową:

   a) $y' = 2x + 1$, $y(-1) = -2$, b) $y' = 3x^2$, $y(0) = -1$,

c) $y'=e^{-x}$, $y(0)=2$, d) $y'=\frac{1}{x}$, $y(-1)=2$,

e) $y'=\frac{\cdot}{x^2}$, $y(1)=-1$, f) $y'=\frac{-2}{x^3}$, $y(-2)=1$.

9. Sprawdzić, że funkcja $y=\frac{1}{1+x^2}$, $x\in R$, jest rozwiązaniem równania $2y''-3(1+x^2)(y')^2+4y^3=0$.

10. Sprawdzić, że funkcja $y=xe^{-x^2}$, $x\in R$, jest rozwiązaniem równania $y''+2xy'+4y=0$.

11. Rozwiązać równanie $y''=2$ i naszkicować dwie dowolnie wybrane krzywe całkowe tego równania.

12. Rozwiązać równanie $y''=\sin x$ przy warunkach początkowych $y(0)=2$, $y'(0)=-1$ oraz naszkicować odpowiednią krzywą całkową.

13. Znaleźć rozwiązanie równania $y''+\frac{1}{x^2}=0$ spełniające warunki brzegowe $y(1)=2$, $y(e)=3$, a następnie naszkicować odpowiednią krzywą całkową.

14. Znaleźć wszystkie rozwiązania danego równania spełniające warunek $y(x_0)=y_0$:

a) $y''=\sin x$, $y(0)=0$, b) $y''=\frac{1}{x}$, $y(-1)=1$.

15. Znaleźć wszystkie funkcje $y=y(x)$, $x\in R_+$, dla których

a) $y''=2$, $y(1)=1$, b) $y''=\frac{2}{x^3}$, $y(1)=0$.

Naszkicować wykresy trzech spośród tych funkcji.

16. Rozwiązać równanie, a następnie znaleźć rozwiązanie szczególne spełniające podane warunki początkowe:

a) $y''=3\cos x$, $y(0)=-3$, $y'(0)=1$,

b) $x^2y''-2x^2+4=0$, $y(1)=-1$, $y'(1)=4$,

c) $y'''=6+e^{-x}$, $y(0)=0$, $y'(0)=1$, $y''(0)=-1$.

17. Rozwiązać równanie, a następnie znaleźć to rozwiązanie, które spełnia podane warunki brzegowe:

a) $y''=2-\sin x$, $y(0)=0$, $y(\pi)=\pi$,

b) $x^2y''-2x^2+4=0$, $y(-1)=1-e$, $y(-e)=4$,

c) $y'''=\frac{6}{x^4}$, $y(-1)=0$, $y(-2)=-1/2$, $y(-4)=-3/4$.

Odpowiedzi. Występujące w odpowiedziach C, C_1, C_2, C_3 oznaczają dowolne stałe.

4. Wsk. Zastosować tw. 1.2 Cauchy'ego o istnieniu i jednoznaczności rozwiązań równania $y'=f(x,y)$.

7. a) $y=x^2-x+C$, b) $y=x^3+C$, c) $y=e^x+C$, d) $y=\sin 2x+C$,

e) $y=\ln|x+1|+C$, f) $y=\frac{-1}{x}+C$, g) $y=-\frac{1}{2x^2}+C$, h) $y=\sqrt{x}+C$, $x>0$.

Sporządzenie odpowiednich rysunków pozostawiamy Czytelnikowi.

8. a) $y=x^2+x-2$, b) $y=x^3-1$, c) $y=-e^{-x}+3$, d) $y=\ln(-x)+2$,

e) $y=-\frac{3}{x}+2$, $x>0$, f) $y=\frac{1}{x^2}+\frac{3}{4}$, $x<0$. Sprządzenie wykresów pozostawiamy Czytelnikowi.

14. a) Rozwiązanie ogólne: $y=C_2+C_1x-\sin x$. Uwzględniając warunek $y(0)=0$ otrzymujemy $C_2=0$. Podany warunek spełnia więc nieskończenie wiele rozwiązań danego równania i są one określone wzorem $y=C_1x-\sin x$, $C_1\in R$.

b) $y=x\ln(-x)-x+Cx+C$.

15. a) $y=x^2+Cx-C$ $x>0$, b) $y=\frac{1}{x}+Cx-C-1$, $x>0$. Sprządzenie wykresów pozostawiamy Czytelnikowi.

16. a) $y=-3\cos x+C_1x+C_2$; $y=x-3\cos x$,

b) $y=x^2+4\ln|x|+C_1x+C_2$; $y=x^2-2x+4\ln x$,

c) $y=x^3-e^{-x}+\frac{1}{2}C_1x^2+C_2x+C_3$; $y=x^3-e^{-x}+1$.

17. a) $y=x^2+\sin x+C_1x+C_2$; $y=x^2+(1-\pi)x+\sin x$,

b) $y=x^2+4\ln|x|+C_1x+C_2$; $y=x^2+4\ln(-x)+ex$,

c) $y=-\frac{1}{x}+\frac{1}{2}C_1x^2+C_2x+C_3$; $y=-\frac{1}{x}-1$, $x<0$.

2. RÓWNANIE O ZMIENNYCH ROZDZIELONYCH. RÓWNANIE JEDNORODNE.

RÓWNANIE O ZMIENNYCH ROZDZIELONYCH. Równanie różniczkowe pierwszego rzędu postaci

(2.1) $$\frac{dy}{dx} = f(x)g(y),$$

gdzie f i g są funkcjami ciągłymi na pewnych przedziałach (w szczególności mogą to być przedziały niewłaściwe), a $y = y(x)$ jest funkcją niewiadomą, nazywamy **równaniem o zmiennych rozdzielonych**.

Na przykład

$$y' = 2xy, \quad y' = x^2(y+1), \quad y' = x-2, \quad y' = y^2 - 1$$

są takimi równaniami.

TWIERDZENIE 2.1 (**o istnieniu i jednoznaczności rozwiązań równania o zmiennych rozdzielonych**). Jeżeli funkcja f jest ciągła na przedziale (a,b), a funkcja g jest ciągła i różna od zera na przedziale (c,d), to dla dowolnej pary liczb (x_0,y_0) takiej, że $x_0 \in (a,b)$, $y_0 \in (c,d)$ istnieje dokładnie jedno rozwiązanie $y = y(x)$ równania (2.1) spełniające warunek $y(x_0) = y_0$ i jest ono określone wzorem

(2.2) $$y = G^{-1}(F(x) - F(x_0) + G(y_0)),$$

gdzie F i G są dowolnie wybranymi funkcjami pierwotnymi funkcji f i $\frac{1}{g}$.

Geometrycznie to oznacza, że *przez każdy punkt obszaru*

$$D = \{(x,y) \in R^2: \ a < x < b \ \wedge \ c < y < d\}$$

przechodzi wtedy dokładnie jedna krzywa całkowa równania (2.1).

D o w ó d . Niech (x_0,y_0) będzie dowolnie ustalonym punktem obszaru D. Przypuśćmy, że $y = y(x)$, $x \in (\alpha,\beta) \subset (a,b)$ jest rozwiązaniem równania (2.1) spełniającym warunek początkowy $y(x_0) = y_0$. Wówczas dla $x \in (\alpha,\beta)$ mamy:

(1) $$y'(x) = f(x)g(y(x)),$$

czyli

(2) $$\frac{y'(x)}{g(y(x))} = f(x),$$

a następnie

(3) $$\int_{x_0}^{x} \frac{y'(x)}{g(y(x))}dx = \int_{x_0}^{x} f(x)dx,$$

(4) $$\int_{y_0}^{y} \frac{dy}{g(y)} = \int_{x_0}^{x} f(x)dx,$$

(5) $$G(y) - G(y_0) = F(x) - F(x_0),$$

gdzie F i G są dowolnie wybranymi funkcjami pierwotnymi funkcji f i $\frac{1}{g}$.

Z założenia twierdzenia i równości $G'(y) = \frac{1}{g(y)}$ wynika, że pochodna $G'(y)$ ma stały znak na przedziale (c,d). Funkcja $G(y)$ jest więc ściśle monotoniczna, a zatem istnieje dokładnie jedna funkcja do niej odwrotna G^{-1}. W konsekwencji z równości (5) otrzymujemy wzór

$$y = G^{-1}(F(x) - F(x_0) + G(y_0)),$$

który jednoznacznie określa rozwiązanie $y = y(x)$.

I odwrotnie: dla dowolnego punktu $(x_0,y_0) \in D$ funkcja $y = y(x)$ określona powyższym wzorem spełnia kolejno równości od (5) do (1), a to oznacza, że jest rozwiązaniem równania (2.1), co kończy dowód. □

Z powyższego wynika, że $y = y(x)$ jest rozwiązaniem równania (2.1) na pewnym przedziale wtedy i tylko wtedy, gdy

(2.3) $$G(y) = F(x) + C,$$

przy czym stała C przyjmuje wartości z pewnego przedziału (stała C nie zawsze może przyjmować dowolne wartości - por. przykład 2.2). Jeżeli potrafimy rozwiązać efektywnie równanie (2.3) względem y, to otrzymamy wzór

(2.4) $$y = G^{-1}(F(x) + C)$$

określający rozwiązanie ogólne równania (2.1). Wzór (2.4) zawiera wszystkie rozwiązania tego równania w obszarze D.

U w a g a . Równanie o zmiennych rozdzielonych może posiadać ponadto pewne rozwiązania szczególne przebiegające poza obszarami, w których spełnione są założenia twierdzenia 2.1. A mianowicie: jeżeli $\hat{y}$ jest taką liczbą, dla której $g(\hat{y})=0$, to funkcja stała $y=\hat{y}$, $x\in(a,b)$, jest rozwiązaniem równania (2.1). Rozwiązanie to może być regularne lub nie (por. przykłady 2.1 i 2.5).

P R Z Y K Ł A D 2.1. Rozwiążemy równanie

(1) $$\frac{dy}{dx}=2(y-1).$$

Jest to równanie o zmiennych rozdzielonych. Przyjmijmy, że $f(x)=2$, $g(y)=y-1$. Funkcja f jest ciągła dla $x\in R$, funkcja g jest ciągła i różna od zera dla $y\neq 1$. Z twierdzenia 2.1 wynika, że przez każdy punkt obszarów

$$D_1=\{(x,y)\in R^2:\ -\infty<x<+\infty\ \wedge\ -\infty<y<1\}$$

oraz

$$D_2=\{(x,y)\in R^2:\ -\infty<x<+\infty\ \wedge\ 1<y<+\infty\}$$

przechodzi dokładnie jedna krzywa całkowa równania (1).

Równanie (2.3) określające wszystkie rozwiązania w obszarach D_1 i D_2 uzyskujemy przez "rozdzielenie zmiennych" w równaniu (1) i całkowanie:

$$\frac{dy}{y-1}=2dx,\quad \int\frac{dy}{y-1}=2\int dx,$$

$$\ln|y-1|=2x+\tilde{C},$$

gdzie $\tilde{C}$ jest dowolną stałą. Przyjmując $\tilde{C}=\ln C$, $C>0$, otrzymujemy dalej

$$\ln|y-1|=\ln e^{2x}+\ln C,$$

$$|y-1|=Ce^{2x},\quad C>0.$$

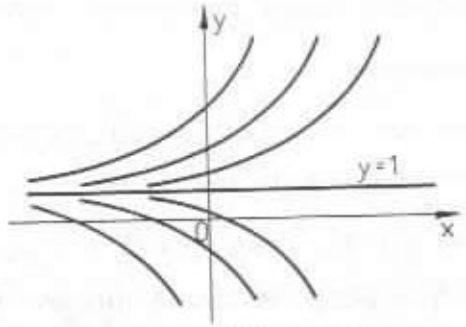
Stąd wynika, że rozwiązania w obszarze D_1 określone są wzorem

$$y=1-Ce^{2x},\quad x\in(-\infty,+\infty),\ C>0,$$

a rozwiązania w obszarze D_2 mają postać

$$y=1+Ce^{2x},\quad x\in(-\infty,+\infty),\ C>0.$$

Zauważmy ponadto, że $g(y)=0$ dla $y=1$. Funkcja $y=1$, $x\in R$, jest także rozwiązaniem równania (1). Dodajmy, że jest to rozwiązanie regularne, co wynika z twierdzenia 1.2 (Cauchy'ego), gdyż prawa strona równania (1) i jej pochodna względem y są funkcjami ciągłymi na całej płaszczyźnie R^2.



Rys 2.1.

Reasumując, przez każdy punkt płaszczyzny przechodzi dokładnie jedno rozwiązanie równania (1), przy czym wszystkie te rozwiązania

$$y=1-Ce^{2x},\ C>0;\quad y=1+Ce^{2x},\ C>0;\quad y=1$$

można zapisać jednym wzorem

$$y=1+\hat{C}e^{2x},\quad x\in R,\ \hat{C}\in R.$$

Jest to rozwiązanie ogólne równania (1). Kilka krzywych całkowych tego równania przedstawia rys. 2.1. ■

P R Z Y K Ł A D 2.2. Znajdziemy rozwiązanie ogólne równania

(1) $$\frac{dy}{dx}=-\frac{x}{y}.$$

Równanie rozważamy w obszarach D_1 i D_2:

$$D_1=\{(x,y)\in R^2:\ -\infty<x<+\infty\ \wedge\ -\infty<y<0\},$$

$$D_2=\{(x,y)\in R^2:\ -\infty<x<+\infty\ \wedge\ 0<y<+\infty\}.$$

Wiadomo, zgodnie z twierdzeniem 2.1, że przez każdy punkt tych obszarów przechodzi dokładnie jedno rozwiązanie. Rozwiązania te znajdujemy przez rozdzielenie zmiennych i całkowanie (przed całkowaniem wygodnie jest pomnożyć obie strony przez 2):

$$y\,dy = -x\,dx, \quad \int 2y\,dy = -\int 2x\,dx,$$

(2) $$y^2 = -x^2 + C, \quad C > 0.$$

Zauważmy, że stała C nie może przyjmować tu dowolnych wartości - nie może być ani zerem, ani liczbą ujemną. Mówimy, że wartościami dopuszczalnymi stałej C są w tym przypadku tylko wartości dodatnie.

Z (2) otrzymujemy

$$y = -\sqrt{C - x^2}, \quad -\sqrt{C} < x < \sqrt{C}, \quad C > 0,$$

gdy rozwiązania przebiegają w obszarze D_1 oraz

$$y = \sqrt{C - x^2}, \quad -\sqrt{C} < x < \sqrt{C}, \quad C > 0,$$

gdy rozwiązania przebiegają w obszarze D_2. Krzywe całkowe są połówkami okręgów $x^2 + y^2 = C$, $C > 0$ leżącymi odpowiednio w półpłaszczyznach D_1 lub D_2. ■

PRZYKŁAD 2.3. Rozwiążemy równanie

(1) $$\frac{dy}{dx} = \frac{2xy}{x^2 - 1}$$

przy warunku początkowym $y(0) = 2$.

Z twierdzenia 2.1 wynika, że przez dowolny punkt każdego obszaru, w którym $x \neq 1$, $x \neq -1$ i $y \neq 0$, przechodzi dokładnie jedna krzywa całkowa. Rozwiązania te znajdujemy przez rozdzielenie zmiennych w równaniu (1) i całkowanie. Kolejno mamy:

$$\int \frac{dy}{y} = \int \frac{2x}{x^2 - 1}\,dx$$

$$\ln|y| = \ln|x^2 - 1| + \ln \widetilde{C}, \quad \widetilde{C} > 0,$$

$$|y| = \widetilde{C}|x^2 - 1|, \quad \widetilde{C} > 0.$$

Ponadto rozwiązaniem równania (1) na każdym z przedziałów: $(-\infty, -1)$, $(-1,1)$ oraz $(1, +\infty)$ jest funkcja stała $y = 0$, przy czym z twierdzenia 1.2 (Cauchy'ego) wynika regularność tego rozwiązania.

Wszystkie rozwiązania można zapisać jednym wzorem

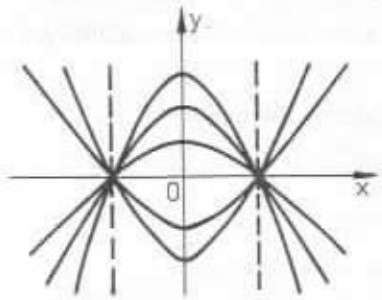
(2) $$y = C(x^2 - 1), \quad x \neq \pm 1, \quad C \in R.$$

Wzór ten określa rozwiązanie ogólne równania (1), a rysunek 2.2 przedstawia kilka krzywych całkowych tego równania.

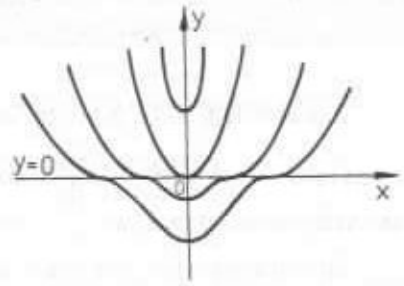
Pozostaje znalezienie rozwiązania szczególnego spełniającego warunek początkowy $y(0) = 2$. Podstawiając $x = 0$ i $y = 2$ w (2) otrzymujemy $C = -2$. Zatem

$$y = 2(1 - x^2), \quad x \in (-1,1)$$

jest szukanym rozwiązaniem szczególnym. ■



Rys 2.2.



Rys 2.3.

PRZYKŁAD 2.4. Znajdziemy rozwiązanie ogólne równania

(1) $$\frac{dy}{dx} = 2\sqrt{y},$$

a następnie rozwiązanie szczególne spełniające warunek $y(4) = 1$.

Jest to równanie o zmiennych rozdzielonych. Z twierdzenia 2.1 wynika, że przez każdy punkt półpłaszczyzny

$$D = \{(x,y) \in R^2: -\infty < x < +\infty \wedge 0 < y < +\infty\}$$

przechodzi dokładnie jedna krzywa całkowa tego równania. Ponadto rozwiązaniem równania (1) jest funkcja stała $y = 0$, $x \in R$.

Aby znaleźć rozwiązania przebiegające w obszarze D, zakładamy, że $y > 0$, rozdzielamy zmienne i całkujemy:

$$\int \frac{dy}{2\sqrt{y}} = \int dx,$$

$$\sqrt{y}=x+C,\quad C\in R,\quad x+C>0.$$

Stąd otrzymujemy rozwiązanie ogólne równania (1):

(2) $$y=(x+C)^2,\quad C\in R,\quad x+C>0.$$

Przyjmując w (2) $x=4$ i $y=1$ otrzymujemy

$$1=(4+C)^2,\quad 4+C>0,$$

skąd wynika, że $C=-3$. Rozwiązanie szczególne spełniające warunek $y(4)=1$ określone jest więc wzorem

$$y=(x-3)^2,\quad x>3.$$

Czytelnikowi pozostawiamy naszkicowanie kilku krzywych całkowych równania (1), w tym krzywej całkowej przechodzącej przez punkt $(4,1)$. ■

PRZYKŁAD 2.5. Rozwiążemy równanie

(1) $$\frac{dy}{dx}=6x\sqrt[3]{y^2}.$$

Przez każdy punkt obszarów

$$D_1=\{(x,y)\in R^2:\ -\infty<x<+\infty\ \wedge\ 0<y<+\infty\},$$

$$D_2=\{(x,y)\in R^2:\ -\infty<x<+\infty\ \wedge\ -\infty<y<0\}$$

przechodzi dokładnie jedna krzywa całkowa tego równania (tw.2.1). Rozdzielając zmienne w równaniu (1) i całkując otrzymujemy:

$$\int\frac{1}{3}y^{-\frac{2}{3}}dy=\int 2x dx,$$

$$y^{\frac{1}{3}}=x^2+C,$$

(2) $$y=(x^2+C)^3,$$

gdzie C jest dowolną stałą. Wzór ten określa rozwiązanie ogólne równania (1). Ponadto funkcja $y=0$, $x\in R$, jest rozwiązaniem szczególnym tego równania, przy czym nietrudno sprawdzić, że jest to rozwiązanie osobliwe. Rysunek 2.3 przedstawia krzywe całkowe tego równania.

Z dokładniejszej analizy rozwiązania ogólnego (2) wynika, że:

a) rozwiązania, których wykresy przebiegają w obszarze D_1 są określone wzorami

$$y=(x^2+C)^3,\quad C>0,\quad x\in(-\infty,+\infty),$$

$$y=x^6,\quad x\in(-\infty,0);\quad y=x^6,\quad x\in(0,+\infty),$$

$$y=(x^2+C)^3,\quad C<0,\quad x\in(-\infty,-\sqrt{-C}),$$

$$y=(x^2+C)^3,\quad C<0,\quad x\in(\sqrt{-C},+\infty);$$

b) rozwiązania, których wykresy przebiegają w obszarze D_2 są określone wzorami

$$y=(x^2+C)^3,\quad C<0,\quad x\in(-\sqrt{-C},\sqrt{-C}).$$ ■

Uwaga. Na ogół pomijać będziemy szczegółową analizę rozwiązań równania. Celem naszym będzie uzyskanie wspólnego wzoru określającego wszystkie lub prawie wszystkie rozwiązania, czyli znalezienie rozwiązania ogólnego danego równania różniczkowego.

PRZYKŁAD 2.6. Rozwiążemy równanie

(1) $$\frac{dy}{dx}=2x(y^2-y).$$

Zakładamy, że $x\in R$, $y\neq 0$ i $y\neq 1$, rozdzielamy zmienne i całkujemy:

$$\int\frac{dy}{y^2-y}=\int 2x dx,$$

$$\int\left(\frac{1}{y-1}-\frac{1}{y}\right)dy=\int 2x dx,$$

$$\ln\left|\frac{y-1}{y}\right|=x^2+\ln\widetilde{C},\quad \widetilde{C}>0,$$

$$\left|\frac{y-1}{y}\right|=\widetilde{C}e^{x^2},\quad \widetilde{C}>0,$$

Ponadto funkcje $y=0$ i $y=1$ dla $x\in R$ są rozwiązaniami szczególnymi równania (1). Zatem wzory

$$\left|\frac{y-1}{y}\right|=\widetilde{C}e^{x^2},\quad \widetilde{C}>0;\quad y=0;\quad y=1$$

określają wszystkie rozwiązania tego równania. Można je zapisać w postaci:

$$y=\frac{1}{1-Ce^{x^2}},\quad C\in R,\quad y=0$$

(rozwiązanie $y=1$ otrzymujemy z pierwszego wzoru przy $C=0$). ■

PRZYKŁAD 2.7. Znajdziemy krzywe $y=y(x)$, dla których odcinek normalnej w dowolnym punkcie, zawarty między punktem krzywej i osią 0x ma stałą długość równą a, $a>0$.

Piszemy równanie normalnej w dowolnym punkcie (x_0,y_0) krzywej $y=y(x)$:

$$y=y_0-\frac{1}{y'(x_0)}(x-x_0).$$

Punkt Q przecięcia normalnej z osią 0x wyznaczymy z układu równań

$$\begin{cases}y=y_0-(x-x_0)/y'(x_0),\\ y=0.\end{cases}$$

Punkt Q ma współrzędne $(x_0+y_0y'(x_0),0)$. Ponieważ długość odcinka $\overline{PQ}$ normalnej jest równa a, więc

$$(y_0y'(x_0))^2+y_0^2=a^2.$$

Z warunków zadania wynika więc, że dowolny punkt (x,y) szukanych krzywych spełnia równanie

(1) $$y^2(y')^2+y^2=a^2.$$

Jest to równanie różniczkowe tych krzywych. Rozwiązując równanie (1) względem y' otrzymujemy

(2) $$y'=-\sqrt{a^2-y^2}/y$$

lub

(3) $$y'=\sqrt{a^2-y^2}/y.$$

Rozwiązujemy najpierw równanie (2). Funkcje stałe $y=a$ i $y=-a$ dla $x\in R$ są rozwiązaniami szczególnymi tego równania. Przy założeniu, że $y\in(-a,0)$ lub $y\in(0,a)$ znajdujemy pozostałe rozwiązania równania poprzez rozdzielenie zmiennych i całkowanie:

$$\int\frac{-y}{\sqrt{a^2-y^2}}dy=\int dx,$$

$$\sqrt{a^2-y^2}=x+C,$$

czyli

$$(x+C)^2+y^2=a^2,$$

przy założeniu, że $x+C>0$.

Analogicznie otrzymujemy, że rozwiązanie ogólne równania (3) ma postać

$$-\sqrt{a^2-y^2}=x+C,$$

czyli

$$(x+C)^2+y^2=a^2,$$

przy założeniu, że $x+C<0$.

Równanie (1) ma więc rozwiązanie określone wzorami: $y=a$ i $y=-a$ oraz $(x+C)^2+y^2=a^2$ przy założeniu, że $y\neq 0$. Krzywymi spełniającymi żądany warunek są: proste $y=a$ i $y=-a$ oraz każdy okrąg o środku w dowolnym punkcie osi 0x i promieniu równym a, z wyłączeniem punktów przecięcia tego okręgu z osią 0x. ■

W następnym przykładzie wyznaczymy trajektorie ortogonalne pewnej rodziny krzywych. Wyjaśnijmy krótko to pojęcie:

Niech będzie dana na płaszczyźnie jednoparametrowa rodzina krzywych.

Trajektoria izogonalna jednoparametrowej rodziny krzywych to krzywa, która w każdym punkcie tworzy stały kąt φ z przechodzącą przez ten punkt krzywą tej rodziny.

W szczególnym przypadku, gdy $\varphi=\pi/2$, trajektorię nazywamy **ortogonalną**.

Rozważmy najprostszy przykład: niech będzie dana rodzina prostych równoległych do prostej $y=x$. Trajektorią ortogonalną tej rodziny jest prosta $y=-x$, a także każda prosta równoległa do niej, czyli $y=-x+C$, gdzie C jest dowolną stałą. Otrzymaliśmy więc całą rodzinę trajektorii ortogonalnych.

Z definicji trajektorii ortogonalnych wynika metoda otrzymania równania różniczkowego tych trajektorii, gdy dane jest równanie różniczkowe rozważanej rodziny krzywych:

$$F(x,y,y')=0.$$

A mianowicie: wystarczy w powyższym równaniu zastąpić y' przez $-1/y'$. W ten sposób otrzymane równanie

$$F(x,y,-1/y')=0$$

jest równaniem różniczkowym trajektorii ortogonalnych.

PRZYKŁAD 2.8. Wyznaczymy trajektorie ortogonalne rodziny parabol o równaniach

(1) $$y^2=2ax,$$

gdzie a jest dowolnym, różnym od zera parametrem.

Najpierw napiszemy równanie różniczkowe tej rodziny krzywych. W tym celu różniczkujemy (y traktujemy jako funkcję x) obie strony równania (1) i otrzymujemy

(2) $$2yy'=2a,$$

a następnie z równań (1) i (2) rugujemy parametr a. W ten sposób otrzymujemy równanie $y^2=2xyy'$, czyli

(3) $$y=2xy'$$

(rozwiązanie $y=0$ możemy pominąć, gdyż prosta ta nie należy do rozważanej rodziny krzywych). Jest to równanie różniczkowe danej rodziny krzywych. Zastępując w równaniu (3) y' przez $-1/y'$ otrzymujemy równanie różniczkowe trajektorii ortogonalnych $y=-2x/y'$, czyli

(4) $$y'=-\frac{2x}{y}.$$

Jest to równanie o zmiennych rozdzielonych, zatem metodę całkowania tego równania już znamy. Otrzymujemy kolejno:

$$\int 2y\,dy=-\int 4x\,dx,$$

$$y^2=-2x^2+C,\quad C>0$$

a stąd

$$2x^2+y^2=C,\quad C>0,$$

$$\frac{x^2}{C/2}+\frac{y^2}{C}=1,\quad C>0.$$

Trajektoriami ortogonalnymi rozważanej rodziny parabol są elipsy symetryczne względem obu osi układu, o ogniskach położonych na osi 0y i osiach $2\sqrt{C}$ i $2\sqrt{C/2}$, przy czym C oznacza dowolną stałą dodatnią. ■

RÓWNANIE JEDNORODNE względem x i y jest to równanie postaci

(2.5) $$\frac{dy}{dx}=f\left(\frac{y}{x}\right),$$

gdzie funkcja $f(t)$ jest funkcją ciągłą swego argumentu na pewnym przedziale, a $y=y(x)$ jest funkcją niewiadomą.

Na przykład równania

$$\frac{dy}{dx}=\frac{y}{x}+1,\quad \frac{dy}{dx}=\frac{y}{x}\left(1+\ln\frac{y}{x}\right),\quad \frac{dy}{dx}=\frac{x^2+y^2}{xy}$$

są równaniami jednorodnymi względem x i y.

Równanie jednorodne sprowadzamy do równania o zmiennych rozdzielonych stosując podstawienie

$$\frac{y}{x}=t(x),$$

czyli

(2.6) $$y=tx,$$

gdzie t oznacza nową niewiadomą funkcję zmiennej x. Różniczkując obie strony (2.6) względem x, otrzymujemy

$$\frac{dy}{dx}=\frac{dt}{dx}x+t.$$

Uwzględniając tę równość w (2.5) mamy

$$\frac{dt}{dx}x+t=f(t),$$

a następnie

(2.7) $$\frac{dt}{dx}=\frac{f(t)-t}{x}.$$

Równanie to jest równaniem o zmiennych rozdzielonych, w którym niewiadomą jest funkcja t zmiennej x.

Sposób znalezienia rozwiązania równania postaci (2.5) pokażemy w przykładach.

PRZYKŁAD 2.9. Rozwiążemy równanie

$$\frac{dy}{dx}=\frac{2y-x}{x}, \tag{1}$$

przy warunku początkowym $y(-1)=2$.

Jest to równanie jednorodne względem x i y, gdyż może być zapisane w postaci

$$\frac{dy}{dx}=2\frac{y}{x}-1.$$

Stosujemy podstawienie $y=tx$. Wówczas

$$\frac{dy}{dx}=t+x\frac{dt}{dx}.$$

Po podstawieniu do równania (1) otrzymujemy

$$t+x\frac{dt}{dx}=2t-1,$$

czyli

$$\frac{dt}{dx}=\frac{t-1}{x}. \tag{2}$$

Równanie (2) jest równaniem o zmiennych rozdzielonych. Funkcja $t=1$ dla $x\in(-\infty,0)$ oraz dla $x\in(0,+\infty)$ jest rozwiązaniem tego równania. Zakładamy, że $t\neq 1$ i szukamy pozostałych rozwiązań:

$$\int\frac{dt}{t-1}=\int\frac{dx}{x},$$

$$\ln|t-1|=\ln|x|+\ln\widetilde{C},\quad \widetilde{C}>0,$$

$$|t-1|=\widetilde{C}|x|,\quad \widetilde{C}>0.$$

Dołączając do ostatniego wzoru rozwiązanie szczególne $t=1$, wszystkie rozwiązania równania (2) zapisujemy jednym wzorem

$$t=1+Cx,\quad C\in R,\quad x\neq 0.$$

Wzór ten określa rozwiązanie ogólne równania (2). Ponieważ $y=tx$, więc rozwiązanie ogólne równania (1) ma postać

$$y=(1+Cx)x, \tag{3}$$

gdzie C oznacza dowolną stałą, $x\in(-\infty,0)$ lub $x\in(0,+\infty)$. Pozostaje znalezienie rozwiązania szczególnego, które spełnia warunek $y(-1)=2$.

Podstawiając $x=-1$ i $y=2$ w (3) otrzymujemy $C=3$. Zatem szukane rozwiązanie szczególne jest określone wzorem

$$y=(1+3x)x,\quad x<0. \tag{4}$$

Zauważmy jeszcze, że z twierdzenia 1.2 (Cauchy'ego) wynika, że przez każdy punkt półpłaszczyzn

$$D_1=\{(x,y)\in R^2:\ x<0\},\qquad D_2=\{(x,y)\in R^2:\ x>0\}$$

przechodzi dokładnie jedna krzywa całkowa równania (1).

Czytelnikowi pozostawiamy naszkicowanie kilku krzywych całkowych równania (2) i (1), a także krzywej określonej wzorem (4). ■

PRZYKŁAD 2.10. Rozwiążemy równanie

$$\frac{dy}{dx}=\frac{4y(x^2+y^2)}{x(3y^2+2x^2)}. \tag{1}$$

Równanie to rozważamy w dwóch półpłaszczyznach

$$D_1=\{(x,y)\in R^2:\ x<0\},\qquad D_2=\{(x,y)\in R^2:\ x>0\},$$

przy czym z twierdzenia 1.2 (Cauchy'ego) wynika, że przez każdy punkt tych obszarów przechodzi dokładnie jedna krzywa całkowa.

Równanie (1) można zapisać w postaci

$$\frac{dy}{dx}=\frac{4(\frac{y}{x}+(\frac{y}{x})^3)}{3(\frac{y}{x})^2+2},$$

więc jest to równanie jednorodne względem x i y. Stosując podstawienie $y=tx$ sprowadzamy równanie (1) do równania o zmiennych rozdzielonych:

$$\frac{dt}{dx}=\frac{t^3+2t}{x(3t^2+2)} \tag{2}$$

z nową niewiadomą funkcją $t=t(x)$.

Funkcja $t=0$ dla $x\in(-\infty,0)$ oraz $x\in(0,+\infty)$ jest rozwiązaniem szczególnym równania (2). Zakładając następnie, że $t\neq 0$ rozdzielamy zmienne i całkujemy:

$$\int\frac{(3t^2+2)dt}{t^3+2t}=\int\frac{dx}{x},$$

skąd otrzymujemy

$$\ln|t^3+2t| = \ln|x| + \ln\widetilde{C},\quad \widetilde{C}>0,$$

$$|t^3+2t| = \widetilde{C}|x|,\quad \widetilde{C}>0.$$

Wzory

$$|t^3+2t| = \widetilde{C}|x|,\quad \widetilde{C}>0;\qquad t=0$$

można zapisać w postaci

(3) $$t^3+2t = Cx,\quad C\in R,$$

przy czym $x\in(-\infty,0)$ lub $x\in(0,+\infty)$. Wzór (3) określa rozwiązanie ogólne równania (2) w postaci uwikłanej.

Ponieważ $y=tx$, więc z (3) otrzymujemy rozwiązanie ogólne równania (1) również w postaci uwikłanej:

$$y^3+2yx^2-Cx^4=0,\quad C\in R,$$

przy czym $x\in(-\infty,0)$ lub $x\in(0,+\infty)$. ■

ZADANIA DO ROZWIĄZANIA.

1. Podać trzy przykłady równań o zmiennych rozdzielonych.
2. Podać przykład równania o zmiennych rozdzielonych, którego rozwiązaniami są funkcje $y=0$ i $y=1$ dla $x\in R$.
3. Znaleźć wszystkie rozwiązania i naszkicować kilka krzywych całkowych równania:

a) $y'+2xy^2=0$, b) $y'=2y\operatorname{ctg}2x$, c) $y'+2\sqrt{y}=0$,

d) $y'+2y=0$, e) $y'+y^2=0$, f) $y'-4x\sqrt{y}=0$,

g) $y'=\dfrac{2y}{x-1}$, h) $y'=\dfrac{2xy}{x^2-4}$, i) $y'+\dfrac{2y}{x\ln x}=0$.

4. Wykazać, że dowolne rozwiązanie $y=y(x)$ równania $y'=(1-2x)y$ spełnia warunki:

$$\lim_{x\to-\infty} y(x)=0 \quad\text{i}\quad \lim_{x\to+\infty} y(x)=0.$$

5. Wykazać, że dla dowolnego rozwiązania $y=y(x)$ równania $y'=\dfrac{x^2+2}{x^3}y$ zachodzi jedna z równości

$$\lim_{x\to0-} y(x)=0 \quad\text{lub}\quad \lim_{x\to0+} y(x)=0.$$

6. Rozwiązać równanie

a) $y'=e^{-y}$, b) $y'=2x(y+2)$, c) $y'=y^2$

d) $y'=\dfrac{-x}{2y^3}$, e) $y'=\dfrac{x+1}{y}$, f) $y'=\dfrac{y^2+y}{x}$,

g) $y'=\dfrac{y\ln y}{x}$, h) $y'=(2x-1)y^2$, i) $2y'=y^2-1$,

j) $y'=\dfrac{\cos x}{y}$, k) $y'=\dfrac{1}{3y^2+1}$, l) $y'=\dfrac{2\sqrt{y}}{x-1}$,

ł) $y'=\dfrac{y^2}{x\sqrt{1-\ln x}}$, m) $y'=\dfrac{y}{2y^2+1}$, n) $y'=\dfrac{2x-1}{e^y}$.

7. Znaleźć rozwiązanie szczególne równania spełniające podany warunek początkowy:

a) $y'=\dfrac{y+1}{x-1}$, $y(0)=2$, b) $y'=\dfrac{1}{y}$, $y(1)=-2$,

c) $y'=\dfrac{2y\ln y}{x}$, $y(1)=e$, d) $y'=y^2-1$, $y(0)=-2$,

e) $y'=y\cos x$, $y(0)=1$, f) $y'=\dfrac{2xy}{1+x^2}$, $y(1)=-2$,

g) $y'=\dfrac{xy^2}{\sqrt{4-x^2}}$, $y(0)=1/2$, h) $y'=2\sqrt{y}\cos x$, $y(\pi/2)=1/4$.

8. Rozwiązać równanie:

a) $y'=\dfrac{x+3y}{x}$, b) $y'=\dfrac{3y^2+2xy}{x^2+2xy}$,

c) $y'=\dfrac{2x+y}{x}$, d) $xy'=y(\ln y-\ln x)$,

e) $y'=\dfrac{y+\sqrt{x^2+y^2}}{x}$, $x>0$, f) $y'=\dfrac{x^2+y^2}{2xy}$.

9. Znaleźć rozwiązanie szczególne równania spełniające podany warunek początkowy:

a) $y' = \frac{2y-x}{x}$, $y(1)=3$, b) $y' = \frac{y+\sqrt{x^2+y^2}}{x}$, $y(-1)=0$,

c) $y' = \frac{2x+y}{x}$, $y(-1)=-2$, d) $xy' = y(\ln y - \ln x)$, $y(1)=e^2$.

Odpowiedzi. Litera C występująca w odpowiedziach oznacza stałą, która przyjmuje, o ile nie ma innej informacji, dowolne wartości rzeczywiste.

3. a) $y=\frac{1}{x^2+C}$, $y=0$, b) $y=C\sin 2x$, $x\neq\frac{k\pi}{2}$, c) $y=(C-x)^2$ dla $x<C$; $y=0$,

d) $y=Ce^{-2x}$, e) $y=\frac{1}{x+C}$, $y=0$, f) $y=(C+x^2)^2$ przy zał. $C+x^2>0$; $y=0$,

g) $y=C(x-1)^2$, $x\neq 1$, h) $y=C(x^2-4)$, $x\neq\pm 2$, i) $y=\frac{C}{\ln^2 x}$. Sporządzenie rysunków pozostawiamy Czytelnikowi.

6. a) $y=\ln(x+C)$, b) $y=-2+Ce^{x^2}$, c) $y=0$, $y=\frac{1}{C-x}$,

d) $y^4=C-x^2$, $-\sqrt{C}<x<\sqrt{C}$, $C>0$, e) $y=\pm\sqrt{x^2+2x+C}$, $y\neq 0$,

f) $y=-1$, $x\neq 0$; $y=\frac{Cx}{1-Cx}$, $x\neq 0$, g) $y=e^{Cx}$, $x\neq 0$, h) $y=0$, $y=\frac{1}{C+x-x^2}$,

i) $y=-1$, $y=\frac{1+Ce^x}{1-Ce^x}$, j) $y^2-2\sin x=C$, $y\neq 0$, k) $y^3+y-x=C$,

l) $y=0$, $x\neq 1$; $y=(C+\ln|x-1|)^2$, przy zał. $C+\ln|x-1|>0$,

ł) $y=0$, $x\in(0,e)$; $y=\frac{1}{2\sqrt{1-\ln x+C}}$, m) $y=0$, $y^2+\ln|y|-x=C$,

n) $y=\ln(x^2-x+C)$.

7. a) $y=2-3x$, $x<1$, b) $y=-\sqrt{2x+2}$, $x>-1$, c) $y=e^{x^2}$, $x>0$,

d) $y=\frac{1+3e^{2x}}{1-3e^{2x}}$, $x>-\ln\sqrt{3}$, e) $y=e^{\sin x}$, $x\in R$, f) $y=-1-x^2$, $x\in R$,

g) $y=\frac{1}{\sqrt{4-x^2}}$, $x\in(-2,2)$, h) $y=(\sin x-\frac{1}{2})^2$, $x\in(\frac{\pi}{6},\frac{5}{6}\pi)$.

8. a) $y=x(Cx^2-1)/2$, $x\neq 0$, b) $y^2+xy-Cx^3=0$, $x\neq 0$, $y\neq -x/2$,

c) $y=Cx+x\ln x^2$, $x\neq 0$, d) $y=xe^{1+Cx}$, $x>0$,

e) $y=(C^2x^2-1)/2C$, $C>0$, $x>0$, f) $x^2-y^2-Cx=0$, $x\neq 0$, $y\neq 0$.

9. a) $y=x(1+2x)$, $x>0$, b) $y=(x^2-1)/2$, $x<0$, c) $y=x(\ln x^2+2)$, $x<0$,

d) $y=xe^{x+1}$, $x>0$.

3. RÓWNANIE LINIOWE I RZĘDU. RÓWNANIE BERNOULLIEGO.

RÓWNANIE LINIOWE. Równanie różniczkowe postaci

$$y' + p(x)y = q(x), \tag{3.1}$$

gdzie p i q są funkcjami ciągłymi na pewnym przedziale, a $y=y(x)$ jest funkcją niewiadomą, nazywamy **równaniem liniowym rzędu pierwszego**.

Jest to równanie liniowe względem niewiadomej funkcji y i jej pochodnej y'.

Gdy $q(x)\equiv 0$, równanie (3.1) ma postać

$$y' + p(x)y = 0 \tag{3.2}$$

i nazywane jest **równaniem liniowym jednorodnym**. Jeżeli natomiast funkcja q nie jest tożsamościowo równa zeru, równanie (3.1) nazywane jest **równaniem liniowym niejednorodnym**.

Równania

$$y' + 3y = x^2, \qquad y' - xy = \sin \qquad y' - y\ln x = 0$$

są równaniami liniowymi, przy czym dwa pierwsze - to równania liniowe niejednorodne, a ostatnie - jednorodne.

Niżej podajemy **twierdzenie o istnieniu i jednoznaczności rozwiązań równania liniowego pierwszego rzędu**.

TWIERDZENIE 3.1. Jeżeli p i q w równaniu (3.1) są funkcjami ciągłymi na przedziale (a,b), to dla dowolnej pary liczb (x_0, y_0), gdzie $x_0\in(a,b)$, istnieje dokładnie jedno rozwiązanie $y=y(x)$ równania (3.1), spełniające warunek początkowy $y(x_0)=y_0$.

Geometrycznie teza tego twierdzenia oznacza, że *przez każdy punkt obszaru*

$$D=\{(x,y)\in R^2:\ a<x<b\ \wedge\ -\infty<y<+\infty\}$$

przechodzi dokładnie jedna krzywa całkowa równania (3.1).

Uwaga. Przedział (a,b) może być przedziałem niewłaściwym.

Dowód. Równanie (3.1) zapiszmy w postaci normalnej

$$y' = -p(x)y + q(x).$$

Zgodnie z założeniem, prawa strona tego równania $f(x,y) = -p(x)y + q(x)$ oraz jej pochodna $f'_y(x,y) = -p(x)$ są funkcjami ciągłymi na obszarze D. Z twierdzenia 1.2 (Cauchy'ego) o istnieniu i jednoznaczności rozwiązań równania $y' = f(x,y)$ wynika natychmiast, że przez każdy punkt obszaru D przechodzi dokładnie jedna krzywa całkowa równania (3.1). □

W dalszym ciągu przyjmujemy, że funkcje p i q są ciągłe na przedziale (a,b), przy czym przedział ten może być niewłaściwy.

ROZWIĄZANIE RÓWNANIA LINIOWEGO JEDNORODNEGO. Zauważmy, że równanie liniowe jednorodne (3.2) jest równaniem o zmiennych rozdzielonych:

$$\frac{dy}{dx} = -p(x)y.$$

Jednym z rozwiązań tego równania jest funkcja $y = 0$, $x \in (a,b)$. Zakładając, że $y \neq 0$, rozdzielamy zmienne i całkujemy

$$\int \frac{dy}{y} = -\int p(x)dx.$$

Oznaczając przez P dowolną z funkcji pierwotnych funkcji p na przedziale (a,b) (tzn., że $P'(x) = p(x)$) mamy:

$$\ln|y| = -P(x) + \ln \widetilde{C}, \quad \widetilde{C} > 0$$

$$|y| = \widetilde{C}e^{-P(x)}, \quad \widetilde{C} > 0.$$

Dołączając do tego rozwiązanie szczególne $y = 0$ otrzymujemy wzór

(3.3) $$y = Ce^{-P(x)}, \quad C \in R, \quad x \in (a,b),$$

który określa rozwiązanie ogólne równania (3.2), przy czym zawiera on wszystkie rozwiązania tego równania.

METODA UZMIENNIANIA STAŁEJ rozwiązywania równań liniowych niejednorodnych. Wykażemy, że każde rozwiązanie równania (3.1) można otrzymać ze wzoru (3.3) zastępując stałą C odpowiednio dobraną funkcją C(x).

Przypuśćmy najpierw, że C(x) jest już tak dobraną funkcją, że

$$y = C(x)e^{-P(x)}, \quad x \in (a,b),$$

jest rozwiązaniem równania (3.1). Zatem

$$C'(x)e^{-P(x)} - C(x)e^{-P(x)}P'(x) + p(x)C(x)e^{-P(x)} = q(x),$$

a stąd

$$C'(x)e^{-P(x)} = q(x),$$

$$C(x) = \int q(x)e^{P(x)}dx,$$

czyli

$$C(x) = Q(x) + C,$$

gdzie C oznacza dowolną stałą, a Q - dowolną z funkcji pierwotnych funkcji $q(x)e^{P(x)}$. W konsekwencji

(3.4) $$y = (Q(x) + C)e^{-P(x)},$$

gdzie C jest dowolną stałą, $P'(x) = p(x)$, $Q'(x) = q(x)e^{P(x)}$ dla $x \in (a,b)$.

I odwrotnie: łatwo sprawdzić (przez podstawienie), że każda funkcja postaci (3.4) jest rozwiązaniem równania (3.1) na przedziale (a,b).

Co więcej, wzór (3.4) zawiera wszystkie rozwiązania równania (3.1), gdyż dla dowolnego punktu $(x_0, y_0) \in D$ można dobrać taką wartość stałej C w (3.4), że rozwiązanie określone tym wzorem spełnia warunek $y(x_0) = y_0$. Istotnie: podstawiając w (3.4) $x = x_0$ i $y = y_0$ otrzymujemy $y_0 = (Q(x_0) + C)e^{-P(x_0)}$, skąd $C = y_0e^{P(x_0)} - Q(x_0)$.

Oznacza to, że wzór (3.4) określa rozwiązanie ogólne równania (3.1), przy czym zawiera on wszystkie rozwiązania tego równania w obszarze D. Dodajmy, że wszystkie te rozwiązania są określone na całym przedziale (a,b).

Metodę uzyskania wzoru (3.4) ze wzoru (3.3) nazywamy metodą uzmienniania stałej.

Zauważmy jeszcze, że rozwiązanie (3.4) można zapisać w postaci sumy $y = y_0(x,C) + y_s(x)$, gdzie

$$y_0 = Ce^{-P(x)}$$

jest rozwiązaniem ogólnym równania (3.2), a

$$y_s = Q(x)e^{-P(x)}$$

jest pewnym rozwiązaniem równania (3.1). Dokładniejsza analiza wzoru (3.4) pozwala sformułować następujące twierdzenie:

TWIERDZENIE 3.2. Jeżeli $y = y_0(x,C)$, $x \in (a,b)$, jest rozwiązaniem ogólnym równania liniowego jednorodnego (3.2) oraz $y = y_s(x)$, $x \in (a,b)$, jest dowolnym rozwiązaniem szczególnym równania niejednorodnego (3.1), to

$$y = y_0(x,C) + y_s(x),$$

jest rozwiązaniem ogólnym równania (3.1).

PRZYKŁAD 3.1. Znajdziemy rozwiązanie ogólne równania

(1) $$y' - \frac{y}{x} = 2\ln x.$$

Z twierdzenia 3.1 wynika, że przez każdy punkt obszaru $D = \{(x,y) \in R^2: x > 0\}$ przechodzi dokładnie jedna krzywa całkowa.

Najpierw rozwiązujemy odpowiednie równanie liniowe jednorodne

$$y' - \frac{y}{x} = 0,$$

czyli

$$\frac{dy}{dx} = \frac{y}{x}.$$

Jest to równanie o zmiennych rozdzielonych, zatem znaną metodą otrzymujemy rozwiązanie ogólne

(2) $$y_0 = Cx, \quad C \in R, \ x > 0,$$

tego równania.

Rozwiązanie ogólne równania (1) znajdziemy "uzmieniając stałą" C we wzorze (2), to znaczy szukać będziemy rozwiązania równania (1) w postaci

$$y = C(x)$$

Obliczamy: $y' = C'(x)x + C(x)$. Podstawiamy do równania (1) i otrzymujemy

$$C'(x)x + C(x) - C(x) = 2\ln x,$$

a stąd

$$C'(x) = \frac{2}{x}\ln x,$$

$$C(x) = \ln^2 x + C, \quad C \in R.$$

Zatem rozwiązanie ogólne równania (1) określa wzór

$$y = (\ln^2 x + C)x, \quad C \in R, \ x > 0. \qquad \blacksquare$$

Rozwiązanie ogólne równania (3.1) można wyznaczyć korzystając z twierdzenia 3.2, co zilustrujemy przykładem.

PRZYKŁAD 3.2. Rozwiążemy równanie

(1) $$y' - y\cos x = \sin 2x,$$

a następnie znajdziemy rozwiązanie tego równania spełniające warunek początkowy $y(0) = -1$.

Najpierw rozwiązujemy równanie jednorodne

$$y' - y\cos x = 0,$$

czyli

(2) $$\frac{dy}{dx} = y\cos x.$$

Rozwiązanie ogólne tego równania ma postać

(3) $$y_0 = Ce^{\sin x}, \quad C \in R, \ x \in R.$$

Następnie szukamy rozwiązania szczególnego równania (1) metodą uzmienniania stałej:

$$y_s = C(x)e^{\sin x}$$

Po zróżniczkowaniu i podstawieniu do równania (1) otrzymujemy

$$C'(x)e^{\sin x} = \sin 2x,$$

$$C'(x) = e^{-\sin x}\sin 2x.$$

Stąd

$$C(x) = \underset{C}{2}(-\sin x - 1)e^{-\sin x}.$$

Zatem funkcja

(4) $$y_s = -2(1 + \sin x), \quad x \in R,$$

jest rozwiązaniem szczególnym równania (1).

Suma rozwiązania ogólnego równania jednorodnego (2) i rozwiązania szczególnego równania niejednorodnego (1), czyli $y = y_0 + y_s$ jest rozwiązaniem ogólnym równania (1). Zatem, zgodnie z (3) i (4), rozwiązanie ogólne równania (1) określa wzór

(5) $y = Ce^{\sin x} - 2(1+\sin x), \quad C \in R, \ x \in R.$

Uwzględniając w (5) warunek $y(0) = -1$ otrzymujemy $C = 1$, zatem funkcja

$$y = e^{\sin x} - 2(1+\sin x), \quad x \in R,$$

jest rozwiązaniem spełniającym podany warunek początkowy. ■

METODA PRZEWIDYWANIA. Jeżeli funkcja p w równaniu (3.1) jest stałą i prawa strona, czyli funkcja q, jest odpowiedniej postaci, to umiemy p r z e w i d z i e ć postać rozwiązania szczególnego. Metoda ta pozwala znaleźć rozwiązanie szczególne równania (3.1) bez uciążliwego czasem całkowania, które występuje przy stosowaniu metody uzmieniania stałej.

Załóżmy, że w równaniu (3.1) $p(x) \equiv a = \text{const}$. Wówczas równanie to przyjmuje postać

(3.5) $$y' + ay = q(x).$$

I. *Jeżeli* $q(x) = W_n(x)e^{\alpha x}$, *gdzie* W_n *oznacza wielomian stopnia n,* $\alpha \in R$, *to istnieje rozwiązanie szczególne równania* (3.5) *określone wzorem*

$$y_s = x^k P_n(x) e^{\alpha x},$$

gdzie P_n *jest pewnym wielomianem stopnia n, natomiast* $k = 1$, *gdy* $\alpha = -a$ *albo* $k = 0$, *gdy* $\alpha \neq -a$.

Tak więc, dla równań

a) $y' + 2y = 3e^x$, b) $y' - y = x^2 + 3x$, c) $y' + y = xe^{-x}$

będziemy szukać rozwiązań postaci

a) $y_s = Ae^x$, b) $y_s = Ax^2 + Bx + C$, c) $y_s = x(Ax + B)e^{-x}$,

gdzie A, B, C są współczynnikami, które należy obliczyć.

P R Z Y K Ł A D 3.3. Znajdziemy rozwiązanie szczególne równania

(1) $$y' + 2y = (4x - 1)e^{-2x}.$$

Przewidujemy, że rozwiązaniem równania (1) jest funkcja

(2) $$y_s = x(Ax + B)e^{-2x}, \quad x \in R,$$

gdzie A i B są odpowiednio dobranymi współczynnikami. Podstawiamy y_s i y_s' do równania (1) i otrzymujemy równość, która winna być spełniona dla każdego $x \in R$:

$$(2Ax + B)e^{-2x} \equiv (4x - 1)e^{-2x}.$$

Stąd wynika, że $A = 2$ i $B = -1$. Zatem

$$y_s = (2x^2 - x)e^{-2x}, \quad x \in R.$$ ■

II. *Jeżeli* $q(x) = W_n(x)\cos\beta x + V_m(x)\sin\beta x$, *gdzie* W_n, V_m *są wielomianami odpowiednich stopni,* $\beta \in R$, *to istnieje rozwiązanie szczególne równania* (3.5) *określone wzorem*

$$y_s = P_l(x)\cos\beta x + Q_l(x)\sin\beta x,$$

gdzie P_l i Q_l *są pewnymi wielomianami stopnia l,* $l = \max(m,n)$.

Zatem dla równań

a) $y' - y = 3\sin x$, b) $y' + y = x\cos 2x - 2\sin 2x$

istnieją rozwiązania postaci

a) $y_s = A\sin x + B\cos x$, b) $y_s = (Ax + B)\cos 2x + (Cx + D)\sin 2x$.

P R Z Y K Ł A D 3.4. Rozwiązanie równania

(1) $$y' + y = 5\cos 2x$$

przewidujemy w postaci

$$y_s = A\cos 2x + B\sin 2x.$$

Po podstawieniu y_s i y_s' do (1) otrzymujemy

$$-2A\sin 2x + 2B\cos 2x + A\cos 2x + B\sin 2x \equiv 5\cos 2x,$$
$$(-2A + B)\sin 2x + (2B + A)\cos 2x \equiv 5\cos 2x.$$

Stąd

$$-2A + B = 0 \quad \text{i} \quad 2B + A = 5,$$

czyli $A = 1$, $B = 2$. Zatem

$$y_s = \cos 2x + 2\sin 2x, \quad x \in R,$$

jest rozwiązaniem szczególnym równania (1). ■

III. *Jeżeli* $q(x) = (W_n(x)\cos\beta x + V_m(x)\sin\beta x)e^{\alpha x}$, *gdzie* W_n, V_m *są wielomianami odpowiednich stopni,* $\alpha \in R$, $\beta \in R$, *to istnieje rozwiązanie szczególne równania* (3.5) *określone wzorem*

$$y_s = x^k(P_l(x)\cos\beta x + Q_l(x)\sin\beta x)e^{\alpha x},$$

gdzie P_l i Q_l *są pewnymi wielomianami stopnia* l, $l=\max(m,n)$ *oraz* $k=1$, *gdy* $\alpha=-a$ *i* $\beta=0$ *albo* $k=0$, *gdy* $\alpha\neq -a$ *lub* $\beta\neq 0$.

Łatwo widać, że przypadek III jest najogólniejszy i z niego wynika I (gdy $\beta=0$) i II (gdy $\alpha=0$).

Powyższe wskazówki nie wystarczą by przewidzieć rozwiązanie szczególne równań:

$$y'-y=1+\sin x, \qquad y'+2y=x-e^x.$$

W takich przypadkach można skorzystać z następującego twierdzenia.

TWIERDZENIE 3.3. Jeżeli $y=y_1(x)$, $x\in(a,b)$, jest rozwiązaniem równania

$$y'+ay=q_1(x),$$

a $y=y_2(x)$, $x\in(a,b)$, jest rozwiązaniem równania

$$y'+ay=q_2(x),$$

to $y=y_1(x)+y_2(x)$, $x\in(a,b)$, jest rozwiązaniem równania

$$y'+ay=q_1(x)+q_2(x).$$

PRZYKŁAD 3.5. Dla równania

(1) $$y'+2y=4x-3e^{-2x}$$

przewidujemy rozwiązanie szczególne

$$y_s=Ax+B+Cxe^{-2x}.$$

Podstawiając y_s i y_s' do równania (1) otrzymujemy równość

$$A+Ce^{-2x}+2Ax+2B\equiv 4x-3e^{-2x},$$

z której wynika, że $A=2$, $B=-1$, $C=-3$. Zatem funkcja

$$y=2x-1-3xe^{-2x}, \quad x\in R,$$

jest rozwiązaniem szczególnym równania (1). ■

PRZYKŁAD 3.6. Rozwiążemy równanie

(1) $$y'-y=2x\cos x$$

przy warunku początkowym $y(0)=2$.

Najpierw rozwiązujemy równanie jednorodne

(2) $$y'-y=0.$$

Rozwiązanie ogólne tego równania określa wzór

(3) $$y_0=Ce^x, \quad C\in R, \quad x\in R.$$

Korzystając z metody przewidywania znajdziemy rozwiązanie szczególne równania (1). Przewidujemy, że

$$y_s=(Ax+B)\cos x+(Cx+D)\sin x.$$

Po podstawieniu y_s i y_s' do równania (1) mamy

$$(-Ax+Cx+A-B+D)\cos x+(-Ax-Cx-B+C-D)\sin x\equiv 2x\cos x.$$

Równość ta zachodzi jedynie wtedy, gdy

$$-Ax+Cx+A-B+D\equiv 2x \quad \text{i} \quad -Ax-Cx-B+C-D\equiv 0,$$

czyli

$$\begin{cases} -A+C &= 2,\\ A-B+D &= 0,\\ -A-C &= 0,\\ -B+C-D &= 0. \end{cases}$$

Stąd otrzymujemy, że $A=-1$, $B=0$, $C=1$, $D=1$. Tak więc, funkcja

(4) $$y_s=-x\cos x+(x+1)\sin x, \quad x\in R,$$

jest rozwiązaniem szczególnym równania (1).

Suma rozwiązania ogólnego równania jednorodnego (2) i rozwiązania szczególnego równania (1) jest rozwiązaniem ogólnym równania (1):

$$y=y_0+y_s.$$

Zatem, zgodnie z (3) i (4), rozwiązanie ogólne równania (1) jest określone wzorem

(5) $$y=Ce^x-x\cos x+(x+1)\sin x, \quad x\in R, \ C\in R.$$

A teraz wyznaczymy to rozwiązanie równania (1), które spełnia warunek $y(0)=2$. Podstawiając $x=0$ i $y=2$ we wzorze (5) otrzymamy $C=2$. Szukanym rozwiązaniem jest funkcja

$$y=2e^x-x\cos x+(x+1)\sin x, \quad x\in R.$$ ■

Czytelnik, który zechce rozwiązać powyższe równanie metodą uzmienniania stałej, może się przekonać, że w tym przypadku metoda przewidywania jest bardziej efektywna, gdyż szybciej i łatwiej prowadzi do rozwiązania zadania.

Niemniej warto pamiętać, że metoda uzmienniania stałej pozostaje podstawową metodą całkowania równania liniowego, gdyż może być stosowana do każdego równania tej postaci (czego nie można powiedzieć o metodzie przewidywania).

RÓWNANIE BERNOULLIEGO. Równanie różniczkowe postaci

$$y' + p(x)y = q(x)y^{\alpha}, \quad \alpha \neq 0, \ \alpha \neq 1, \tag{3.6}$$

gdzie $y = y(x)$ jest funkcją niewiadomą, p i q są funkcjami ciągłymi na pewnym przedziale, przy czym funkcja q nie jest tożsamościowo równa zeru, nazywamy **równaniem Bernoulliego.**

Zauważmy, że dla $\alpha = 0$ równanie (3.6) jest równaniem liniowym niejednorodnym, a dla $\alpha = 1$ równaniem liniowym jednorodnym.

Jeżeli równanie (3.6) zapiszemy w postaci normalnej

$$y' = -p(x)y + q(x)y^{\alpha}$$

i zastosujemy do niego twierdzenie 1.2 (Cauchy'ego) o istnieniu i jednoznaczności rozwiązań równania różniczkowego $y' = f(x,y)$, to łatwo wywnioskujemy co następuje:

Jeżeli w równaniu Bernoulliego funkcje p *i* q *są ciągłe na przedziale* (a,b), *to dla dowolnej pary liczb* (x_0, y_0), *gdzie* $x_0 \in (a,b)$, $y_0 \in (0,+\infty)$, *istnieje dokładnie jedno rozwiązanie* $y = y(x)$ *tego równania spełniające warunek* $y(x_0) = y_0$.

Geometrycznie to oznacza, że *przez każdy punkt obszaru*

$$D_1 = \{(x,y) \in R^2 : a < x < b \wedge y > 0\}$$

przechodzi wtedy dokładnie jedna krzywa całkowa równania (3.6).

Jeżeli α jest taką liczbą, że potęga y^{α} jest określona także dla $y < 0$, to również przez każdy punkt obszaru

$$D_2 = \{(x,y) \in R^2 : a < x < b \wedge y < 0\}$$

przechodzi dokładnie jedna krzywa całkowa równania (3.6)

Zauważmy jeszcze, że w przypadku gdy $\alpha > 0$, funkcja stała $y = 0$, $x \in (a,b)$, jest także rozwiązaniem równania (3.6). Wykazuje się, że jest to rozwiązanie regularne dla $\alpha > 1$, a osobliwe dla $0 < \alpha < 1$.

Aby znaleźć rozwiązania równania Bernoulliego sprowadzamy je do równania liniowego wprowadzając nową niewiadomą funkcję $t = t(x)$ za pomocą równości:

$$y^{1-\alpha}(x) = t(x). \tag{3.7}$$

Wówczas

$$(1-\alpha)y^{-\alpha}\frac{dy}{dx} = \frac{dt}{dx}. \tag{3.8}$$

Dzieląc przez y^{α} obie strony równania (3.6) mamy

$$y^{-\alpha}\frac{dy}{dx} + p(x)y^{1-\alpha} = q(x),$$

a po uwzględnieniu związków (3.7) i (3.8) otrzymujemy równanie liniowe

$$\frac{dt}{dx} + (1-\alpha)p(x)t = (1-\alpha)q(x),$$

w którym $t = t(x)$ jest niewiadomą funkcją.

PRZYKŁAD 3.7. Znajdziemy rozwiązanie ogólne równania

a) $y' - y + e^{-2x}y^2 = 0$, b) $y' - \frac{2y}{x} = 2x\sqrt{y}\cos x$.

a) Jest to równanie Bernoulliego ($\alpha = 2$):

$$y' - y = -e^{-2x}y^2. \tag{1}$$

Z twierdzenia 1.2 (Cauchy'ego) wynika, że przez każdy punkt płaszczyzny przechodzi dokładnie jedna krzywa tego równania. Funkcja $y = 0$, $x \in R$, jest rozwiązaniem szczególnym równania (1). Zakładając, że $y \neq 0$ dzielimy obie strony równania (1) przez y^2

$$\frac{1}{y^2}y' - \frac{1}{y} = -e^{-2x}$$

i stosujemy podstawienie $\frac{1}{y} = t$. Wówczas $-\frac{1}{y^2}y' = t'$. W ten sposób otrzymujemy równanie liniowe z niewiadomą funkcją $t = t(x)$:

$$t' + t = e^{-2x}. \tag{2}$$

Wzór

$$t = Ce^{-x} - e^{-2x}, \quad C \in R,$$

określa rozwiązanie ogólne równania (2). W konsekwencji, wobec związku $1/y=t$, otrzymujemy, że

$$y=\frac{1}{Ce^{-x}-e^{-2x}}, \quad C\in R,$$

jest rozwiązaniem ogólnym równania (1). Ponadto rozwiązaniem tego równania jest funkcja stała $y=0$ dla $x\in R$.

b) Jest to równanie Bernoulliego ($\alpha=1/2$):

(1) $$y'-\frac{2y}{x}=2x\sqrt{y}\cos x.$$

Z twierdzenia 1.2 (Cauchy'ego) wynika, że przez każdy punkt obszarów

$$D_1=\{(x,y)\in R^2:\ x>0\ \wedge\ y>0\},\quad D_2=\{(x,y)\in R^2:\ x<0\ \wedge\ y>0\}$$

przechodzi dokładnie jedna krzywa całkowa tego równania. Ponadto rozwiązaniem równania (1) jest funkcja stała $y=0$ dla $x\in(-\infty,0)$ i dla $x\in(0,+\infty)$. Znajdziemy teraz rozwiązania przebiegające w obszarach D_1 i D_2.

Przy założeniu, że $x\neq 0$ i $y>0$ dzielimy obie strony równania (1) przez $\sqrt{y}$:

$$\frac{1}{\sqrt{y}}y'-\frac{2\sqrt{y}}{x}=2x\cos x,$$

a następnie stosujemy podstawienie $\sqrt{y}=t$. Wówczas $\frac{1}{2\sqrt{y}}y'=t'$. W konsekwencji otrzymujemy równanie liniowe

(2) $$t'-\frac{t}{x}=x\cos x.$$

Stosując metodę uzmienniania stałej otrzymujemy, że

$$t=x\sin x+Cx,\quad C\in R,$$

jest rozwiązaniem ogólnym równania (2).

Ponieważ $t=\sqrt{y}$, więc $y=t^2$ przy założeniu, że $t>0$. Zatem wzór

(3) $$y=(x\sin x+Cx)^2,\ C\in R,\ \text{przy zał., że } x\sin x+Cx>0,$$

określa rozwiązanie ogólne równania (1).

Na koniec wyznaczymy to rozwiązanie równania (1), które przechodzi przez punkt $(\pi/2,\ \pi^2/16)\in D_1$. Podstawiając $x=\pi/2$, $y=\pi^2/16$ w (3) otrzymujemy

$$\frac{\pi^2}{16}=(\frac{\pi}{2}+C\frac{\pi}{2})^2 \quad \text{i} \quad \frac{\pi}{2}+C\frac{\pi}{2}>0,$$

czyli

$$\frac{1}{4}=(1+C)^2 \quad \text{i} \quad 1+C>0.$$

Stąd wynika, że $C=-\frac{1}{2}$.

Rozwiązanie, którego szukamy jest określone na przedziale (a,b) takim, że $(a,b)\subset(0,+\infty)$, $\frac{\pi}{2}\in(a,b)$ oraz, zgodnie z (3), $\sin x-\frac{1}{2}>0$ dla $x\in(a,b)$. Jest to przedział $(\frac{\pi}{6},\frac{5\pi}{6})$. Zatem rozwiązaniem spełniającym podany warunek początkowy jest funkcja

$$y=(x\sin x-\frac{1}{2}x)^2,\quad x\in(\frac{\pi}{6},\frac{5\pi}{6}).$$ ■

ZADANIA DO ROZWIĄZANIA.

1. Podać dwa przykłady równań liniowych 1 rzędu.
2. Podać przykład równania liniowego takiego, że

   a) przez dowolny punkt płaszczyzny przechodzi krzywa całkowa tego równania,

   b) wszystkie krzywe całkowe tego równania przebiegają w półpłaszczyźnie $D=\{(x,y)\in R^2:\ x>0\}$.
3. Rozwiązać następujące równanie liniowe jednorodne:

   a) $y'+2xy=0$, b) $y'-\frac{2xy}{1+x^2}=0$, c) $y'+\frac{y}{x\ln x}=0$,

   d) $y'=y\operatorname{ctg}x$, e) $y'-3y=0$, f) $y'+ay=0$, $a=\text{const}$.
4. Znaleźć rozwiązanie równania spełniające warunek $y(x_0)=y_0$, a następnie naszkicować odpowiednią krzywą całkową:

a) $y'-2y=0,\ y(0)=3$, b) $y'+y=0,\ y(-1)=-2e$,

c) $y'=\frac{y}{x\ln x},\ y(e)=-1$, d) $y'+\frac{y}{x\ln x}=0,\ y(e)=2$,

e) $y'+\frac{2xy}{1+x^2}=0,\ y(0)=-1$, f) $y'=\frac{2x-3}{x^2-3x}y,\ y(1)=2$.

5. Stosując metodę uzmienniania stałej rozwiązać równanie:

a) $y'+y=\frac{e^{-x}}{x^2}$, b) $y'-\frac{2y}{x}=e^{-\frac{1}{x}}$, c) $y'-\frac{2xy}{1+x^2}=2x$,

d) $y'+2xy=2x^3$, e) $y'-\frac{y}{x}=\ln x$, f) $y'+2y=5\sin x$,

g) $y'-y=2e^x\cos^2 x$, h) $x^2y'-2xy=3$, i) $y'+y\operatorname{tg}x=\sin 2x$.

6. Stosując metodę przewidywania znaleźć jedno z rozwiązań równania

a) $y'-y=2+2x-x^2$, b) $y'+y=2e^{-3x}$, c) $y'+2y=e^{2x}$,

d) $y'-2y=xe^x$, e) $y'+y=2e^{-x}$, f) $y'+y=4xe^{-x}$,

g) $y'-3y=(2x+1)e^{3x}$, h) $y'-2y=5\sin x$, i) $y'-y+2x\sin x=0$,

j) $y'+y=3\sin x+\cos x$, k) $y'+2y=6-4\sin 2x$,

l) $y'+2y=e^{-2x}-5\sin x$, l) $y'+y=x^2+2\sin x$.

7. Stosując metodę przewidywania rozwiązać równanie:

a) $y'+y=x^2$, b) $y'-2y=3e^{-x}$, c) $y'+y=(1+2x)e^x$,

d) $y'-y-2xe^x=0$, e) $y'-3y=2e^{3x}$, f) $y'-y=2\cos x$,

g) $y'-2y=(4x-1)e^{2x}$, h) $y'-y=\cos 2x+2\sin 2x$,

i) $y'+y=2x-e^{-x}$, j) $y'+y=\sin x+x(\sin x+\cos x)$,

k) $y'+2y=2x^2-1+4\cos 2x$, l) $y'=y+4e^{-x}+\cos x-\sin x$.

8. Rozwiązać równanie:

a) $y'+y=2\cos x$, b) $y'-\frac{y}{x}=\frac{1}{\ln^2 x}$, c) $y'-\frac{y}{\operatorname{tg}x}=2\sin x$,

d) $y'-2y=(3x+2)e^{-x}$, e) $y'=2xy+e^{x^2}\cos x$, f) $y'-y=-5\sin 2x$,

g) $y'+\frac{y}{x}=\frac{2}{x^4}$, h) $y'-\frac{y}{x\ln x}=\frac{\ln x}{x^2}$, i) $y'=\frac{y}{x}+2x^4e^{-x^2}$,

j) $y'+y=\frac{1}{\sqrt{3+e^x}}$, k) $y'=\frac{y}{2\sqrt{x}}-1$, l) $y'-y=\frac{e^x}{\sqrt{9+x^2}}$.

9. Rozwiązać równanie przy podanym warunku początkowym:

a) $y'+\frac{e^xy}{2+e^x}=x,\ y(0)=-1$, b) $y'-2y=\frac{3x^2-2x}{e^x},\ y(0)=0$,

c) $y'+2y=5\sin x,\ y(0)=3$, d) $y'-y\operatorname{tg}x=e^{\sin x},\ y(0)=1$,

e) $y'+\frac{y}{x\ln x}=\frac{2x}{\ln x},\ y(e)=e^2$, f) $y'+\frac{2xy}{1+x^2}=\frac{1}{x^2},\ y(-1)=-1$,

g) $y'=\frac{y}{1+e^x}+e^x,\ y(0)=2$, h) $y'-y\cos x=\sin 2x,\ y(\pi)=-1$,

i) $y'+y=\frac{2}{(1+e^x)^2},\ y(0)=-1$, j) $y'-\frac{y}{x}-\frac{1}{\ln^2 x}=0,\ y(\frac{1}{e})=0$,

k) $y'=y+x+2\cos x,\ y(0)=1$, l) $y'-2y=e^{2x}+5\cos x,\ y(0)=-2$,

10. Znaleźć rozwiązanie $y=y(x)$ równania $xy'-y=\frac{2x}{\ln^3 x}$ spełniające warunek $y(e)=0$, a następnie obliczyć $\lim\limits_{x\to 1^+}y(x)$ i $\lim\limits_{x\to+\infty}y(x)$.

11. Podać przykład równania Bernoulliego, dla którego jednym z rozwiązań jest funkcja stała $y=0$ dla $x\in R$.

12. Rozwiązać równanie:

a) $y'+y=xy^2$, b) $y'+\frac{y}{2x}=(x^2-x)y^3$, c) $y'+\frac{y}{2x}=\frac{1}{2yx^2}$,

d) $y'=y+\frac{\cos x}{2ye^{-2x}}$, e) $y'+2y=4\sqrt{y}\sin x$, f) $y'+\frac{y}{x}=2y^2\ln x$,

g) $y'+\frac{2y}{x}=4x^2y\sqrt{y}$, h) $y'+\frac{y}{3\operatorname{ctg}x}=\frac{\sin 2x}{y^2}$, i) $y'+\frac{2y}{3x^2}=\frac{2e^{\frac{1}{x}}}{\sqrt{y}}$.

13. Rozwiązać równanie przy podanym warunku początkowym:

a) $y'+\frac{y}{2x}=\frac{1}{y}$, $y(-1)=-2$, b) $y'+\frac{y}{x}=\frac{y^2}{\sqrt{\ln x}}$, $y(e)=\frac{-1}{e}$,

c) $y'-y=\frac{2+e^{2x}}{2y}$, $y(0)=1$, d) $y'=y\operatorname{ctg}x-\frac{y^2}{\sqrt{4-\cos x}}$, $y(\frac{\pi}{2})=1$,

e) $y'+\frac{y}{x\ln x}=\frac{2y^2}{x\ln^2 x}$, $y(e)=\frac{1}{2}$, f) $y'-\frac{2}{3}xy+\frac{2x^3}{y^2}=0$, $y(0)=1$,

g) $y'-2y=2e^x\sqrt{y}$, $y(1)=9e^2$, h) $y'+2y=\frac{4\sqrt{y}\,e^{-x}}{1+4x^2}$, $y(0)=1$.

Odpowiedzi. Występująca w odpowiedziach litera C oznacza dowolną stałą.

3. a) $y=Ce^{-x^2}$, b) $y=C(1+x^2)$, c) $y=C/\ln x$, d) $y=C\sin x$, $x\neq k\pi$,
e) $y=Ce^{3x}$, f) $y=Ce^{-ax}$.

4. a) $y=3e^{2x}$, b) $y=-2e^{-x}$, c) $y=-\ln x$, $x>1$, d) $y=\frac{2}{\ln x}$, $x>1$,
e) $y=\frac{-1}{1+x^2}$, f) $y=3x-x^2$, $0<x<3$. Sporządzenie odpowiednich rysunków pozostawiamy Czytelnikowi.

5. a) $y=(C-\frac{1}{x})e^{-x}$, b) $y=x^2(C+e^{-1/x})$, c) $y=(1+x^2)(C+\ln(1+x^2))$,
d) $y=x^2-1+Ce^{-x^2}$, e) $y=x(C+\frac{1}{2}\ln^2 x)$, f) $y=Ce^{-2x}+2\sin x-\cos x$,
g) $y=(x+\frac{1}{2}\sin 2x+C)e^x$, h) $y=Cx^2-\frac{1}{x}$, i) $y=C\cos x-2\cos^2 x$, $x\neq\frac{\pi}{2}+k\pi$.

6. a) $y=x^2-2$, b) $y=-e^{-3x}$, c) $y=\frac{1}{4}e^{2x}$, d) $y=(-x-1)e^x$,
e) $y=2xe^{-x}$, f) $y=2x^2e^{-x}$, g) $y=(x^2+x)e^{3x}$, h) $y=-\cos x-2\sin x$,
i) $y=(x+1)\cos x+x\sin x$, j) $y=2\sin x-\cos x$, k) $y=3+\cos 2x-\sin 2x$,
l) $y=xe^{-2x}+\cos x-2\sin x$, l) $y=x^2-2x+2+\sin x-\cos x$.

7. a) $y=Ce^{-x}+x^2-2x+2$, b) $y=Ce^{2x}-e^{-x}$, c) $y=Ce^{-x}+xe^x$,
d) $y=(C+x^2)e^x$, e) $y=(C+2x)e^{3x}$, f) $y=Ce^x+\sin x-\cos x$,
g) $y=Ce^{2x}+(2x^2-x)e^{2x}$, h) $y=Ce^x-\cos 2x$, i) $y=Ce^{-x}+2x-2-xe^{-x}$,
j) $y=Ce^{-x}+x\sin x$, k) $y=Ce^{-2x}+x^2-x+\sin 2x+\cos 2x$,
l) $y=Ce^x-2e^{-x}+\sin x$.

8. a) $y=Ce^{-x}+\cos x+\sin x$, b) $y=x(C-\frac{1}{\ln x})$, c) $y=2x\sin x+C\sin x$, $x\neq\frac{k\pi}{2}$,

d) $y=Ce^{2x}-(x+1)e^{-x}$, e) $y=(C+\sin x)e^{x^2}$, f) $y=Ce^x+2\cos 2x+\sin 2x$,
g) $y=\frac{1}{x}(C-\frac{1}{x^2})$, h) $y=(C-\frac{1}{x})\ln x$, $x>0$, i $x\neq 1$,
i) $y=Cx-x(x^2+1)e^{-x^2}$, $x\neq 0$, j) $y=(C+2\sqrt{3+e^x})e^{-x}$,
k) $y=Ce^{\sqrt{x}}+2(1+\sqrt{x})$, $x>0$, l) $y=e^x(C+\ln(x+\sqrt{9+x^2}))$.

9. a) $y=\frac{x^2+xe^x}{2+e^x}-1$, b) $y=-x^2e^{-x}$, c) $y=4e^{-2x}+2\sin x-\cos x$,
d) $y=\frac{1}{\cos x}e^{\sin x}$, $x\in(-\frac{\pi}{2},\frac{\pi}{2})$, e) $y=\frac{x^2}{\ln x}$, $x>1$, f) $y=\frac{x^2-2x-1}{x(1+x^2)}$, $x<0$,
g) $y=\frac{e^x+x+3}{1+e^{-x}}$, h) $y=-2-2\sin x+e^{\sin x}$, i) $y=\frac{-2}{e^{2x}+e^x}$,
j) $y=-x-\frac{x}{\ln x}$, $0<x<1$, k) $y=3e^x-x-1+\sin x-\cos x$,
l) $y=xe^{2x}+\sin x-2\cos x$.

10. $y=x(1-\frac{1}{\ln^2 x})$, $x>1$; $-\infty$ przy $x\to 1+$, $+\infty$ przy $x\to+\infty$.

12. a) $y=0$, $y=\frac{1}{Ce^x+x+1}$, b) $y=0$, $x\neq 0$; $y=\frac{\pm 1}{\sqrt{x(C+2x-x^2)}}$,
c) $y=\pm\sqrt{\frac{1}{x}(C+\ln|x|)}$, $y\neq 0$, d) $y=\pm\sqrt{C+\sin x}\,e^x$, $y\neq 0$,
e) $y=0$, $y=(Ce^{-x}+\sin x-\cos x)^2$, przy zał., że $Ce^{-x}+\sin x-\cos x>0$
f) $y=0$, $x>0$; $y=\frac{1}{x(C-\ln^2 x)}$, g) $y=0$, $x\neq 0$; $y=\frac{1}{x^2(C-x^2)^2}$ przy zał., że $x(C-x^2)>0$, h) $y=\sqrt[3]{(C-6\cos x)\cos x}$, $x\neq\frac{k\pi}{2}$, $y\neq 0$,
i) $y=\sqrt[3]{(3x+C)^2e^{\frac{2}{x}}}$, $3x+C>0$.

13. a) $y=-\sqrt{\frac{x^2-5}{x}}$, $x\in(-\sqrt{5},0)$, b) $y=\frac{1}{x(1-2\sqrt{\ln x})}$, $x>\sqrt[4]{e}$,
c) $y=e^x\sqrt{2+x-e^{-2x}}$, d) $y=\frac{\sin x}{2\sqrt{4-\cos x}-3}$, $0<x<\pi$,
e) $y=\frac{\ln x}{1+\ln^2 x}$, $x>1$, f) $y=\sqrt[3]{3+3x^2-2e^{x^2}}$, $y>0$,
g) $y=(2+x)^2e^{2x}$, $x>-2$, h) $y=(1+\operatorname{arctg}2x)^2e^{-2x}$, $x>-\frac{\pi}{8}$.

4. RÓWNANIE ZUPEŁNE. CZYNNIK CAŁKUJĄCY.

RÓWNANIE ZUPEŁNE. Załóżmy, że funkcje $P(x,y)$ i $Q(x,y)$ są funkcjami klasy C^1 na pewnym jednospójnym obszarze D, przy czym funkcja Q jest różna od zera w punktach tego obszaru.

Równanie różniczkowe

$$\frac{dy}{dx}+\frac{P(x,y)}{Q(x,y)}=0,$$

zapisywane na ogół w postaci

(4.1) $$P(x,y)dx+Q(x,y)dy=0,$$

nazywamy **równaniem zupełnym**, gdy lewa strona równania (4.1) jest różniczką zupełną pewnej funkcji na obszarze D.

Oznacza to, że istnieje taka funkcja F, że

$$\frac{\partial F}{\partial x}=P(x,y) \quad \text{i} \quad \frac{\partial F}{\partial y}=Q(x,y) \quad \text{dla} \quad (x,y)\in D.$$

Funkcję F nazywamy **funkcją pierwotną pary funkcji** P i Q lub **funkcją pierwotną różniczki** $P(x,y)dx+Q(x,y)dy$.

Przy założeniu, że funkcje P i Q są klasy C^1 na obszarze jednospójnym D, *warunkiem koniecznym i wystarczającym istnienia funkcji pierwotnej pary funkcji P i Q jest równość:*

(4.2) $$\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}, \quad (x,y)\in D.$$

Jeśli więc funkcje P, Q są klasy C^1, funkcja Q jest różna od zera oraz $P'_y=Q'_x$ na obszarze jednospójnym D, to równanie (4.1) jest równaniem zupełnym na tym obszarze.

Przy przyjętych założeniach, z twierdzenia 1.2 (Cauchy'ego) wynika, że *przez każdy punkt obszaru D przechodzi dokładnie jedna krzywa całkowa równania* (4.1).

Wykazuje się, że dla każdego rozwiązania $y=y(x)$, $x\in I$, równania (4.1) istnieje taka stała C, że $F(x,y(x))=C$ dla $x\in I$, i odwrotnie: każda funkcja $y=y(x)$, dla której $F(x,y(x))=C$, $x\in I$, gdzie C jest dowolną stałą, jest rozwiązaniem równania (4.1). Zatem równanie

(4.3) $$F(x,y)=C,$$

określa rozwiązanie ogólne równania (4.1).

Łatwo sprawdzić, że rozwiązanie szczególne spełniające warunek $y(x_0)=y_0$ otrzymamy z (4.3) przyjmując $C=F(x_0,y_0)$, przy czym wiadomo, że przy przyjętych założeniach równanie

(4.4) $$F(x,y)=F(x_0,y_0)$$

określa dokładnie jedno takie rozwiązanie (por. rozdz. II, 7, tw. 7.1). Niemniej, efektywne rozwiązanie równania (4.4) względem y bywa trudne, a czasem wręcz niemożliwe. Wówczas pozostajemy przy postaci uwikłanej (4.4) rozwiązania szczególnego lub (4.3) - rozwiązania ogólnego.

PRZYKŁAD 4.1. Równanie

(1) $$(2xy-3x^2)dx+(x^2+3y^2+1)dy=0$$

jest równaniem zupełnym na całej płaszczyźnie. Istotnie:

$$P(x,y)=2xy-3x^2, \quad Q(x,y)=x^2+3y^2+1, \quad P'_y=2x, \quad Q'_x=2x,$$

więc $P'_y=Q'_x$ oraz $Q(x,y)\neq 0$ dla $(x,y)\in R^2$. Przez każdy punkt płaszczyzny przechodzi więc dokładnie jedno rozwiązanie równania (1).

Wyznaczymy funkcję F, dla której lewa strona (1) jest różniczką zupełną, czyli taką funkcję F, dla której

$$\begin{cases} F'_x=2xy-3x^2, \\ F'_y=x^2+3y^2+1. \end{cases}$$

Z pierwszego z tych równań wynika, że

$$F(x,y)=\int(2xy-3x^2)dx=x^2y-x^3+\varphi(y).$$

Różniczkując względem y i uwzględniając drugie równanie, mamy

$$F'_y=x^2+\varphi'(y)=x^2+3y^2+1,$$

skąd $\varphi'(y)=3y^2+1$, $\varphi(y)\underset{C}{=}y^3+y$. Zatem

$$F(x,y)\underset{C}{=}x^2y-x^3+y^3+y.$$

Rozwiązanie ogólne równania (1) określa równanie

(2) $$x^2y-x^3+y^3+y=C,$$

gdzie C jest dowolną stałą.

Następnie znajdziemy rozwiązanie szczególne równania (1) spełniające warunek $y(-1)=1$. Podstawiając $x=-1$ i $y=1$ w (2) otrzymujemy $C=4$. Zatem szukanym rozwiązaniem jest funkcja uwikłana $y=y(x)$ określona równaniem

$$x^2y-x^3+y^3+y-4=0,$$

dla której $y(-1)=1$. Taka funkcja, zgodnie z twierdzeniem o istnieniu funkcji uwikłanej, jest tylko jedna. ■

CZYNNIK CAŁKUJĄCY. Jeżeli dla funkcji P i Q w równaniu (4.1) nie jest spełniony warunek (4.2), to równanie to nie jest równaniem zupełnym. W pewnych przypadkach można wyznaczyć taką funkcję $\mu=\mu(x,y)$ klasy C^1, różną od zera na obszarze D, że

$$\mu(x,y)P(x,y)dx+\mu(x,y)Q(x,y)dy=0 \tag{4.5}$$

jest równaniem zupełnym. Funkcję μ nazywamy wtedy **czynnikiem całkującym** równania (4.1).

Czynnik całkujący μ spełnia więc równanie:

$$\frac{\partial}{\partial y}(\mu P)=\frac{\partial}{\partial x}(\mu Q),$$

czyli

$$\frac{\partial\mu}{\partial y}P+\frac{\partial P}{\partial y}\mu=\frac{\partial\mu}{\partial x}Q+\frac{\partial Q}{\partial x}\mu. \tag{4.6}$$

Jeżeli μ jest funkcją tylko j e d n e j z m i e n n e j: x albo y, to rozwiązanie tego równania nie jest trudne. W dalszym ciągu ograniczymy się do tych dwóch przypadków.

I. Przypuśćmy, że istnieje czynnik całkujący zależny tylko od zmiennej x, czyli $\mu=\mu(x)$. Równanie (4.6) możemy zapisać w postaci

$$\mu P'_y=\mu'(x)Q+\mu Q'_x,$$

skąd ($Q\neq 0$ na obszarze D) otrzymujemy

$$\frac{1}{\mu}\mu'(x)=\frac{P'_y-Q'_x}{Q}. \tag{4.7}$$

Jeżeli prawa strona tego równania nie zależy od zmiennej y, to przy oznaczeniu

$$A(x)=\frac{P'_y-Q'_x}{Q}$$

z równania (4.7) wynika, że

$$\mu(x)=e^{\int A(x)dx}.$$

I odwrotnie: łatwo sprawdzić, że funkcja μ określona tym wzorem jest czynnikiem całkującym równania (4.1).

II. Analogicznie, gdy funkcja P jest różna od zera na obszarze D oraz wyrażenie $(Q'_x-P'_y)/P$ nie zależy od zmiennej x, to istnieje czynnik całkujący μ zależny tylko od zmiennej y, przy czym

$$\mu(y)=e^{\int B(y)dy},$$

gdzie

$$B(y)=\frac{Q'_x-P'_y}{P}.$$

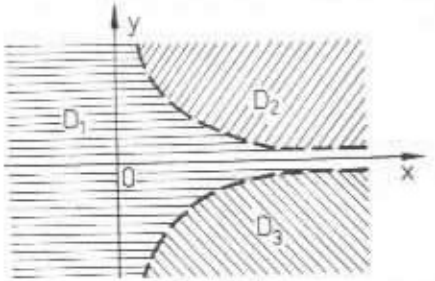
P R Z Y K Ł A D 4.2. Równanie

$$y^2(x-y)dx+(1-xy^2)dy=0 \tag{1}$$

nie jest równaniem zupełnym. Zapiszmy to równanie w postaci normalnej:

$$\frac{dy}{dx}=\frac{y^2(y-x)}{1-xy^2}.$$

Prawa strona tego równania $f(x,y)=\frac{y^2(y-x)}{1-xy^2}$ i jej pochodna f'_y są funkcjami ciągłymi na każdym obszarze, w którym $1-xy^2\neq 0$ (rys 4.1).



Rys 4.1.

Z twierdzenia 1.2 (Cauchy'ego) wynika, że przez każdy punkt tych obszarów przechodzi dokładnie jedno rozwiązanie równania (1). Inaczej mówiąc, w obszarach D_1, D_2 i D_3 (rys. 4.1) ma miejsce jednoznaczność rozwiązań równania (1).

Aby znaleźć rozwiązania tego równania szukamy czynnika całkującego. Ponieważ wyrażenie

$$\frac{P'_y - Q'_x}{Q} = \frac{2y(x-y)}{1-xy^2}$$

zależy od obu zmiennych, więc czynnik całkujący postaci $\mu = \mu(x)$ nie istnieje. Natomiast wyrażenie

$$\frac{Q'_x - P'_y}{P} = -\frac{2}{y} = B(y)$$

nie zależy od x, zatem istnieje czynnik całkujący równania (1) zależny tylko od zmiennej y, przy czym

$$\mu(y) = e^{\int B(y)dy} = e^{-\ln y^2},$$

czyli

$$\mu(y) = \frac{1}{y^2}.$$

Mnożąc obie strony równania (1) przez czynnik całkujący otrzymujemy

$$(2) \qquad (x-y)dx + (\frac{1}{y^2} - x)dy = 0.$$

Jest to równanie zupełne na każdym jednospójnym obszarze, w którym $y \neq 0$ i $1-xy^2 \neq 0$ (por. rys. 4.1).

Wyznaczamy, znanym już sposobem, funkcję pierwotną lewej strony równania (2)

$$F(x,y) \underset{C}{=} \frac{x^2}{2} - xy - \frac{1}{y}.$$

Rozwiązanie ogólne równania (2) i równania (1) otrzymujemy w postaci uwikłanej. Jest ono określone równaniem

$$(3) \qquad \frac{x^2}{2} - xy - \frac{1}{y} = C, \quad C \in R.$$

Ponadto rozwiązaniem równania (1) jest funkcja stała $y = 0$, $x \in R$. Zatem wszystkie rozwiązania $y = y(x)$ danego równania są określone równaniami:

$$\frac{x^2}{2} - xy - \frac{1}{y} = C, \quad C \in R; \qquad y = 0, \quad x \in R.$$

Znajdziemy teraz rozwiązanie szczególne równania (1) spełniające warunek początkowy $y(-2) = 1$. Podstawiając $x = -2$ i $y = 1$ w (3), otrzymujemy $C = 3$. Szukanym rozwiązaniem szczególnym jest funkcja uwikłana $y = y(x)$ określona równaniem

$$\frac{x^2}{2} - xy - \frac{1}{y} - 3 = 0$$

i spełniająca warunek $y(-2) = 1$. ■

ZADANIA DO ROZWIĄZANIA.

1. Znaleźć rozwiązanie ogólne oraz naszkicować kilka krzywych całkowych równania:

   a) $\frac{x}{\sqrt{1+x^2+y^2}}dx + (1 + \frac{y}{\sqrt{1+x^2+y^2}})dy = 0,$

   b) $(2x + \ln(4-y))dx - \frac{x}{4-y}dy = 0.$

2. Znaleźć rozwiązanie ogólne równania:

   a) $(2x - y\cos x + y)dx + (x - \sin x)dy = 0,$

   b) $(2xy - 3x^2)dx + (x^2 + 3y^2 + 1)dy = 0,$

   c) $(y^3 + 2x)dx + (3xy^2 - 2y)dy = 0,$

   d) $(3x^2 - 2xy - 4x + y)dx + (2y + x - x^2)dy = 0,$

   e) $(ye^{-x} + 2xe^{-y})dx - (e^{-x} + x^2e^{-y})dy = 0,$

   f) $(2xe^y + 2e^x)dx + (x^2e^y + 3y^2 + 2)dy = 0,$

   g) $(\frac{2}{x}\ln x + 2xy^2)dx + (2x^2y + \frac{1}{y})dy = 0.$

3. Znaleźć to rozwiązanie równania, które spełnia podany warunek początkowy:

   a) $(3x^2 - 2xy^2)dx + (2y - 2x^2y)dy = 0, \quad y(0) = -2,$

   b) $(2x - y + 1)dx + (2y - x - 1)dy = 0, \quad y(0) = 1,$

   c) $\frac{2x}{y^3}dx + \frac{y^2 - 3x^2}{y^4}dy = 0, \quad y(2) = -1,$

d) $(2ye^{2x}-2x)dx+(e^{2x}+2e^{-2y})dy=0,\quad y(1)=0,$

e) $(2x\ln y-\frac{1}{x})dx+(2y+\frac{x^2}{y})dy=0,\quad y(-1)=1,$

f) $(\sin 2x+2xy)dx+(x^2+\cos y+2)dy=0,\quad y(0)=\pi,$

g) $(3x^2+y\cos x)dx+(3y^2+\sin x+3)dy=0,\quad y(1)=0.$

4. Rozwiązać równanie za pomocą czynnika całkującego:

a) $(1-x^2y)dx+x^2(y-x)dy=0,$

b) $(2xy^2-y)dx+(y^2+x+y)dy=0,$

c) $(x-\frac{y^2}{x})dx+2ydy=0,$

d) $(x+y^2)dx+2ydy=0,$

e) $(x^2+3y^2)dx-2xydy=0,$

f) $(ye^{x+y}-2x)dx+(e^{x+y}+x^2)dy=0,$

g) $(4x^3+\frac{3x^2}{y^2})dx+(\frac{2x^4}{y}+4y)dy=0,$

h) $(y^2\sin x+2xy)dx+(2y-y\cos x+1)dy=0,$

i) $(2\ln(y-x)-\frac{x}{y-x})dx+(\frac{x}{y-x}+\frac{1}{xy})dy=0.$

Odpowiedzi. Litera C w odpowiedziach oznacza stałą, która przyjmuje, jeśli nie ma innej informacji, dowolne wartości rzeczywiste. Czytelnikowi pozostawiamy wyznaczenie obszarów, w których zapewniona jest jednoznaczność rozwiązań odpowiednich równań.

1. a) $y+\sqrt{1+x^2+y^2}=C,\ C>0,$ a stąd $y=\frac{C^2-1-x^2}{2C},\ C>0.$

b) $x^2+x\ln(4-y)=C,\ C\in R,$ a stąd $y=4-e^{\frac{C-x^2}{x}},\ C\in R.$

Sporządzenie odpowiednich rysunków pozostawiamy Czytelnikowi.

2. a) $y=\frac{C-x^2}{x-\sin x},$ b) $x^2y-x^3+y^3+y=C,$ c) $xy^3+x^2-y^2=C,$

d) $x^3-x^2y-2x^2+xy+y^2=C,$ e) $x^2e^{-y}-ye^{-x}=C,$

f) $x^2e^y+2e^x+y^2+2y=C,$ g) $\ln^2 x+x^2y^2+\ln|y|=C.$

3. a) $x^3-x^2y^2+y^2-4=0,\ y(0)=-2,$ czyli $y=-\sqrt{\frac{4-x^3}{1-x^2}},\ x\in(-1,1),$

b) $x^2+y^2-xy+x-y=0,\ y(0)=1,$ czyli $y=\frac{1}{2}(x+1+\sqrt{1-2x-3x^2}),$

c) $\frac{x^2}{y^3}-\frac{1}{y}+3=0,\ y(2)=-1,$ d) $ye^{2x}-e^{-2y}-x^2+2=0,\ y(1)=0,$

e) $x^2\ln y-\ln(-x)+y^2-1=0,\ y(-1)=1,$

f) $\sin^2 x+\sin y+x^2y+2y-2\pi=0,\ y(0)=\pi,$

g) $x^3+y^3+3y+y\sin x-1=0,\ y(1)=0.$

4. a) $\mu(x)=\frac{1}{x^2};\ y^2-2xy-\frac{2}{x}=C,$ b) $\mu(y)=\frac{1}{y^2};\ x^2-\frac{x}{y}+y+\ln|y|=C,\ y=0,$

c) $\mu(x)=\frac{1}{x};\ x+\frac{y^2}{x}=C,$ d) $\mu(x)=e^x;\ y^2=Ce^{-x}-x+1,$

e) $\mu(x)=\frac{1}{x^4};\ \frac{1}{x}+\frac{y^2}{x^3}+C=0,$ f) $\mu(y)=e^{-y};\ ye^x-x^2e^{-y}=C,$

g) $\mu(y)=y^2;\ x^4y^2+x^3+y^4=C,$ h) $\mu(y)=\frac{1}{y};\ x^2+2y-y\cos x+\ln|y|=C,\ y=0,$

i) $\mu(x)=x;\ x^2\ln(y-x)+\ln|y|=C.$

5. RÓWNANIA RÓŻNICZKOWE II RZĘDU SPROWADZALNE DO RÓWNAŃ I RZĘDU.

RÓWNANIE RÓŻNICZKOWE II RZĘDU. Każde równanie różniczkowe II rzędu można zapisać w postaci

$$F(x,y,y',y'')=0, \tag{5.1}$$

gdzie F jest funkcją ciągłą na pewnym obszarze $D\subset R^4$, a $y=y(x)$ oznacza funkcję niewiadomą.

Rozwiązanie ogólne takiego równania jest dwuparametrową rodziną funkcji i można je zapisać jak następuje

$$y=y(x,C_1,C_2).$$

Czasem rozwiązanie ogólne otrzymujemy w postaci uwikłanej, to znaczy jest ono określone równaniem

$$\Phi(x,y,C_1,C_2)=0.$$

PRZYKŁAD 5.1. Rozwiążemy równanie

(1) $$y''=6x$$

przy warunkach początkowych $y(0)=0$, $y'(0)=-1$.

Rozwiązanie ogólne równania (1) znajdujemy przez dwukrotne całkowanie:

$$y'=3x^2+C_1,$$

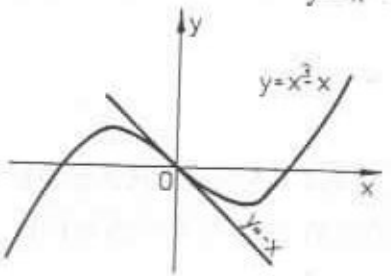
(2) $$y=x^3+C_1x+C_2, \quad x\in R.$$

Wzór (2), w którym C_1 i C_2 oznaczają dowolne stałe, określa rozwiązanie ogólne równania (1). Uwzględniając warunki początkowe mamy

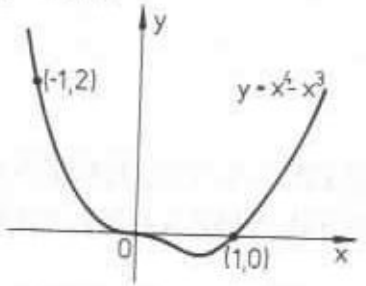
$$y(0)=C_2=0, \qquad y'(0)=C_1=-1.$$

Zatem szukanym rozwiązaniem szczególnym jest funkcja:

$$y=x^3-x, \quad x\in R.$$



Rys 5.1.



Rys 5.2.

Geometrycznie to oznacza, że krzywa całkowa odpowiadająca temu rozwiązaniu (rys. 5.1) przechodzi przez punkt (0,0) i ma w tym punkcie styczną o współczynniku kierunkowym $m=-1$.

Zauważmy na koniec, że warunek $y(0)=0$ spełnia nieskończenie wiele rozwiązań danego równania. Są to rozwiązania, dla których $y(0)=C_2=0$, czyli funkcje postaci:

$$y=x^3+C_1x, \qquad x\in R,\ C_1\in R.$$

Oznacza to, że przez punkt (0,0) przechodzi nieskończenie wiele krzywych całkowych równania (1), ale tylko jedna spośród nich (ta, którą wcześniej wyznaczyliśmy) ma w tym punkcie styczną o podanym kierunku. ■

PRZYKŁAD 5.2. Wyznaczymy rozwiązanie równania

(1) $$y''=12x^2-6x$$

spełniające warunki brzegowe $y(1)=0$, $y(-1)=2$.

Rozwiązanie ogólne tego równania określa wzór

$$y=x^4-x^3+C_1x+C_2, \quad x\in R,$$

gdzie C_1 i C_2 oznaczają dowolne stałe. Uwzględniając warunki brzegowe otrzymujemy

$$y(1)=C_1+C_2=0, \quad y(-1)=2-C_1+C_2=2.$$

Stąd $C_1=0$ i $C_2=0$. Zatem

$$y=x^4-x^3, \quad x\in R,$$

jest rozwiązaniem szczególnym spełniającym podane warunki brzegowe. Krzywa całkowa odpowiadająca temu rozwiązaniu (rys 5.2) przechodzi przez punkty (1,0) i (−1,2). ■

Niżej omówimy rozwiązywanie pewnych równań II rzędu, które przez odpowiednie podstawienie można sprowadzić do równań I rzędu.

RÓWNANIE POSTACI $F(x,y',y'')=0$. W równaniu

(5.2) $$F(x,y',y'')=0$$

nie występuje w sposób wyraźny niewiadoma funkcja y. Stosujemy podstawienie

$$y'=u(x).$$

Wówczas $y''=u'$. Uwzględniając to w równaniu (5.2) otrzymujemy równanie I rzędu:

$$F(x,u,u')=0,$$

w którym $u=u(x)$ jest funkcją niewiadomą.

PRZYKŁAD 5.3. Rozwiążemy równanie

(1) $$y'' + \frac{y'}{x} - 4x^2 = 0.$$

Jest to równanie postaci (5.2) - nie zawiera niewiadomej funkcji y (krótko: równanie " bez y "). Stosujemy podstawienie $y' = u(x)$. Wówczas $y'' = u'$. Po podstawieniu otrzymujemy równanie

(2) $$u' + \frac{u}{x} = 4x^2.$$

Jest to równanie liniowe I rzędu z niewiadomą funkcją $u = u(x)$. Rozwiązanie ogólne równania (2) (wyznaczamy je metodą uzmienniania stałej) określone jest wzorem

$$u = x^3 + \frac{C_1}{x}, \quad C_1 \in R.$$

Ponieważ $u = y'$, pozostaje scałkować równanie

$$y' = x^3 + \frac{C_1}{x}.$$

Stąd otrzymujemy rozwiązanie ogólne równania (1) w postaci

$$y = \frac{1}{4}x^4 + C_1 \ln|x| + C_2,$$

gdzie C_1 i C_2 oznaczają dowolne stałe. ■

RÓWNANIE POSTACI $F(y,y',y'') = 0$. W równaniu postaci

(5.3) $$F(y,y',y'') = 0,$$

w którym niewiadomą jest funkcja $y = y(x)$, **nie występuje w sposób wyraźny zmienna x**. Równanie to sprowadzamy do równania I rzędu stosując podstawienie

$$y' = u(y).$$

Wówczas $y'' = \frac{du}{dy}\frac{dy}{dx} = u\frac{du}{dy}$. Po wykonaniu tego podstawienia w równaniu (5.3) otrzymujemy równanie I rzędu

$$F(y,u,u\frac{du}{dy}) = 0,$$

w którym niewiadomą jest funkcja $u = u(y)$.

PRZYKŁAD 5.4. Rozwiążemy równanie

(1) $$y'' = \frac{(y')^2}{y-1},$$

w którym niewiadomą jest funkcja $y = y(x)$.

Jest to równanie typu (5.3) - nie zawiera zmiennej x (krótko: równanie " bez x "). Stosujemy podstawienie

$$y' = u(y).$$

Wówczas $y'' = u\frac{du}{dy}$. W konsekwencji otrzymujemy równanie I rzędu o zmiennych rozdzielonych

$$u\frac{du}{dy} = \frac{u^2}{y-1}.$$

Rozwiązanie ogólne tego ównania ma postać

$$u = C_1(y-1), \quad C_1 \in R, \quad y \neq 1.$$

Ponieważ $u = y'$, więc

$$y' = C_1(y-1), \quad y \neq 1,$$

a następnie

(2) $$y = 1 + C_2 e^{C_1 x}, \quad C_1 \in R, \quad C \in R \setminus \{0\}.$$

Jest to rozwiązanie ogólne równania (1).

Wyznaczymy teraz rozwiązanie równania (1), które spełnia warunki początkowe: $y(0) = 2$, $y'(0) = -1$. Zgodnie z (2) mamy

$$y = 1 + C_2 e^{C_1 x}, \qquad y' = C_2 C_1 e^{C_1 x},$$

a następnie podstawiając $x = 0$, $y = 2$, $y' = -1$, otrzymujemy

$$2 = 1 + C_2, \qquad -1 = C_2 C_1.$$

Stąd $C_1 = -1$ i $C_2 = 1$. Zatem szukane rozwiązanie określone jest wzorem $y = 1 + e^{-x}$, $x \in R$.

Na koniec znajdziemy rozwiązanie równania (1) spełniajace warunki brzegowe: $y(0) = 0$, $y(1) = 1 - e$.

Uwzględniając te warunki we wzorze (2) mamy

$$0 = 1 + C_2, \quad 1 - e = 1 + C_2 e^{C_1}.$$

Stąd $C_1 = 1$, $C_2 = -1$. Rozwiązanie spełniające podane warunki brzegowe określone jest wzorem $y = 1 - e^x$, $x \in R$. ■

ZADANIA DO ROZWIĄZANIA.

1. Rozwiązać równanie $y'' = 6x$ przy warunkach brzegowych $y(0) = 2$, $y(1) = 1$.
2. Rozwiązać równanie $y'' = (x-2)e^{-x}$ przy warunkach początkowych $y(0) = 1$, $y'(0) = 1$.
3. Znaleźć wszystkie rozwiązania równania $y'' = 6x$, dla których $y(0) = 2$.
4. Rozwiązać równanie $y'' = \frac{2}{x^3}$, a następnie:
   a) znaleźć wszystkie rozwiązania, dla których $y(1) = 1$ i naszkicować wykresy dwóch spośród tych rozwiązań;
   b) znaleźć rozwiązanie, dla którego $y(1) = 1$, $y(4) = -\frac{11}{4}$ i naszkicować wykres tego rozwiązania;
   c) znaleźć rozwiązanie, dla którego $y(1) = 1$, $y'(\frac{1}{2}) = 0$ i naszkicować wykres tego rozwiązania.
5. Sprawdzić, że każda funkcja postaci
   $$y = (1-x)e^x \ln(1-x) + (x+C)e^x, \quad x < 1, \quad C \in R,$$
   jest rozwiązaniem równania $y'' - 2y' + y = \frac{e^x}{1-x}$. Czy jest to rozwiązanie ogólne tego równania w półpłaszczyźnie
   $$D = \{(x,y) \in R^2 : x < 1\}?$$
6. Stosując podstawienie $y' = u(x)$ znaleźć wszystkie rozwiązania równania:
   a) $y'' - \frac{y'}{x} = 0$, b) $y'' + \frac{y'}{x} = \frac{1}{x^3}$, c) $y'' - \frac{y'}{x} = xe^x$,
   d) $y'' + (y')^2 = 0$, e) $y'' \operatorname{tg} x - y' = 1$, f) $y'' = (y')^2 - 1$,

   g) $y'' = \frac{(y')^2 + y'}{x}$, h) $y'' + \frac{y'}{x} = -x(y')^2$, i) $y'' = \frac{y'-1}{2\sqrt{x}}$,
   j) $xy'' = y' \ln y'$, k) $y'' - 2x(y')^2 = 0$, l) $y'' + e^{-x}(y')^2 = 0$.
7. Stosując podstawienie $y' = u(y)$ rozwiązać równanie:
   a) $y^2 y'' = (y')^3$, b) $yy'' - (y')^2 = y^2 y'$, c) $y'(1+y') = y''$,
   d) $yy'' = 2(y')^2$, e) $(3y^2+4)y'' = -6y(y')^2$, f) $y'' - 2(y')^3 = 0$,
   g) $y^2 y'' + y' = 0$, h) $yy'' - (y')^2 + 2y' = 0$, i) $y'' - (y')^2 = 0$,
   j) $(2y - \sin y)y'' = (\cos y - 2)(y')^2$, k) $yy''(2 + \ln y) + (y')^2 = 0$.
8. Rozwiązać równanie przy podanych warunkach początkowych:
   a) $yy'' - (y')^2 = 0$, $y(0) = 1$, $y'(0) = -1$,
   b) $y'' = \frac{y'}{x} + 2\ln x$, $y(1) = \frac{1}{2}$, $y'(1) = \frac{1}{2}$,
   c) $(1+y^2)y'' = 2yy'(1+y')$, $y(-1) = -1$, $y'(-1) = 1$,
   d) $yy'' + (y')^2 = (3y^2 + 2y)y'$, $y(0) = 1$, $y'(0) = 2$,
   e) $1 + y^3 y'' = 0$, $y(2) = -1$, $y'(2) = 1$,
   f) $xy'' = y'(1 - y')$, $y(1) = 1$, $y'(1) = 2$,
   g) $y'' - 2x(y')^2 = 0$, $y(1) = 0$, $y'(1) = -1/2$.
9. Rozwiązać równanie przy podanych warunkach brzegowych:
   a) $y'' - \frac{y'}{x} = \frac{3}{x^3}$, $y(1) = 2$, $y(2) = 9/2$,
   b) $(y+2)y'' - 2(y')^2 = 0$, $y(0) = 0$, $y(1) = -3/2$,
   c) $y'' - \frac{2y'}{x \ln x} = \frac{1}{x}$, $y(e) = 0$, $y(e^2) = -e^2$,
   d) $y'' + (y')^2 = 0$, $y(0) = 0$, $y(1) = 1$,
   e) $yy'' + (y')^2 = 0$, $y(-1) = -4$, $y(1) = -2$,
   f) $y'' = \frac{y'}{2\sqrt{x}}$, $y(1) = 0$, $y(4) = e^2$.

Odpowiedzi. Litery C, C_1, C_2 w odpowiedziach oznaczają stałe, które przyjmują, jeśli nie ma innej informacji, dowolne wartości rzeczywiste.

1. $y = x^3 - 2x + 2$. 2. $y = 1 + xe^{-x}$. 3. $y = x^3 + Cx + 2$.

4. $y = \frac{1}{x} + C_1x + C_2$, a) $y = \frac{1}{x} + Cx - C$, $x > 0$, b) $y = \frac{1}{x} + x - 1$, $x > 0$,
c) $y = \frac{1}{x} + 4x - 4$, $x > 0$. Sporządzenie odpowiednich rysunków pozostawiamy Czytelnikowi

6. a) $y = C_1x^{-1} + C_2$, $x \neq 0$, b) $y = \frac{1}{x} + C_1 \ln|x| + C_2$,
c) $y = (x-1)e^x + C_1x^2 + C_2$, $x \neq 0$,
d) $y = C$; $y = \ln|x + C_1| + C_2$, e) $y = C_2 - C_1\cos x - x$, $x \neq k\pi/2$,
f) $y = x - \ln|1 - C_1e^{2x}| + C_2$, $y = -x + C$,
g) $y = C$, $x \neq 0$; $y = -x + C$, $x \neq 0$; $y = C_2 - x - \frac{1}{C_1}\ln|1 - C_1x|$, $x \neq 0$, $C_1 \neq 0$,
h) $y = C$, $x \neq 0$; $y = -\frac{1}{x} + C$; $y = \frac{1}{C_1}\ln\left|\frac{x}{x + C_1}\right| + C_2$, $C_1 \neq 0$,
i) $y = x + 2C_1(\sqrt{x} - 1)e^{\sqrt{x}} + C_2$, $x > 0$, j) $y = x + C$; $y = \frac{1}{C_1}e^{C_1x} + C_2$, $C_1 \neq 0$,
k) $y = C$; $y = \frac{1}{x} + C$; $y = \frac{1}{2\sqrt{C_1}}\ln\left|\frac{\sqrt{C_1} + x}{\sqrt{C_1} - x}\right| + C_2$ $(C_1 > 0)$,
$y = \frac{-1}{\sqrt{-C_1}}\operatorname{arctg}\frac{x}{\sqrt{-C_1}} + C_2$ $(C_1 < 0)$; albo: $y = \frac{1}{x} + C$, $C \in R$,
oraz $y = C_2 + \frac{1}{2}C_1\ln\left|\frac{1 + C_1x}{1 - C_1x}\right|$, $y = C_2 - C_1\operatorname{arctg}(C_1x)$, gdzie $C_1 \geq 0$, $C_2 \in R$,
l) $y = C$; $y = C - e^x$; $y = \frac{1}{C_1}\ln|C_1e^x - 1| + C_2$, $C_1 \neq 0$.

7. a) $y = C$; $\ln|y| - x + C_1y = C_2$, b) $y = C$; $y = \frac{1}{C - x}$; $y = \frac{C_1C_2e^{C_1x}}{1 - C_2e^{C_1x}}$,
c) $y = C$; $y = -x + C$; $y = -\ln(C_1 - C_2e^x)$, d) $y = C$; $y = \frac{1}{C_2 - C_1x}$, $C_1 \neq 0$,
e) $y = C$; $y^3 + 4y - C_1x = C_2$, f) $y = C$; $y^2 - C_1y + x + C_2 = 0$,
g) $y = C$; $y^2 = 2x + C$; $y - \frac{1}{C_1}\ln|1 + C_1y| - C_1x = C_2$, $C_1 \neq 0$,

h) $y = C$; $y = 2x + C$; $y = \frac{1}{C_1}(C_2e^{C_1x} - 2)$, $C_1 \neq 0$,
i) $y = C$; $y = -\ln(C_2 - C_1x)$, j) $y = C$; $y^2 + \cos y - C_1x = C_2$,
k) $y = C$, $C > 0$; $y(1 + \ln y) - C_1x = C_2$.

8. a) $y = e^{-x}$, b) $y = \frac{x^2}{2}(\ln^2 x - \ln x + 1)$, c) $y = \frac{-1}{x + 2}$, $x > -2$,
d) $y = \frac{e^x}{2 - e^x}$, $x < \ln 2$, e) $y = -\sqrt{5 - 2x}$, $x < 5/2$,
f) $y = x + \ln\sqrt{2x - 1}$, g) $y = \frac{\pi}{4} - \operatorname{arctg} x$.

9. a) $y = \frac{1}{x} + x^2$, $x > 0$, b) $y = \frac{-6x}{1 + 3x}$, $x > -\frac{1}{3}$, c) $y = x(1 - \ln x)$, $x > 1$,
d) $y = \ln((e-1)x + 1)$, e) $y = -\sqrt{10 - 6x}$, $x < 5/3$, f) $y = (\sqrt{x} - 1)e^{\sqrt{x}}$, $x > 0$.

6. RÓWNANIE LINIOWE II RZĘDU.

RÓWNANIE LINIOWE RZĘDU DRUGIEGO. Równanie róż-niczkowe postaci

$$y'' + p_1(x)y' + p_2(x)y = q(x), \tag{6.1}$$

gdzie p_1, p_2 i q są funkcjami ciągłymi na pewnym przedziale, a $y = y(x)$ jest funkcją niewiadomą, nazywamy **równaniem liniowym** II rzędu.

Jest to równanie liniowe względem niewiadomej funkcji y i jej pochodnych y' i y''.

Jeśli $q(x) \equiv 0$, to równanie (6.1) przyjmuje postać

$$y'' + p_1(x)y' + p_2(x) = 0$$

i nazywane jest **równaniem liniowym jednorodnym**. Jeżeli funkcja q nie jest tożsamościowo równa zeru, równanie (6.1) nazywane jest **równaniem liniowym niejednorodnym**.

Na przykład równania

$$y'' - 2xy' = x^3, \qquad y'' - xy' + 2y = 0$$

są równaniami liniowymi rzędu drugiego, przy czym pierwsze z nich jest równaniem liniowym niejednorodnym, a drugie - jednorodnym.

Niżej podajemy **twierdzenie o istnieniu i jednoznaczności rozwiązań równania liniowego drugiego rzędu**:

TWIERDZENIE 6.1. Jeżeli p_1, p_2 i q w równaniu (6.1) są funkcjami ciągłymi na przedziale (a,b), to dla dowolnej trójki liczb (x_0, y_0, y_0'), gdzie $x_0 \in (a,b)$, istnieje dokładnie jedno rozwiązanie $y=y(x)$ równania (6.1) spełniające warunki początkowe $y(x_0)=y_0$, $y'(x_0)=y_0'$.

Geometrycznie teza tego twierdzenia oznacza, że *przez każdy punkt obszaru*

$$D = \{(x,y) \in R^2 : a < x < b \wedge -\infty < y < +\infty\}$$

przechodzi dokładnie jedna krzywa całkowa równania (6.1) *o podanym współczynniku kierunkowym stycznej.*

RÓWNANIE LINIOWE JEDNORODNE. Omówimy teraz całkowanie równania liniowego jednorodnego:

(6.2) $$y'' + p_1(x)y' + p_2(x)y = 0.$$

Najpierw udowodnimy

TWIERDZENIE 6.2. Jeżeli y_1 i y_2 są rozwiązaniami szczególnymi równania (6.2) na przedziale (a,b) oraz C_1 i C_2 są dowolnymi stałymi, to funkcja

$$y = C_1y_1(x) + C_2y_2(x), \qquad x \in (a,b),$$

jest rozwiązaniem równania (6.2).

Krótko: Kombinacja liniowa dowolnych dwóch rozwiązań równania liniowego jednorodnego (6.2) jest również rozwiązaniem tego równania.

D o w ó d . Z założenia wynika, że dla każdego $x \in (a,b)$ zachodzą równości:

$$y_1''(x) + p_1(x)y_1'(x) + p_2(x)y_1(x) = 0,$$

$$y_2''(x) + p_1(x)y_2'(x) + p_2(x)y_2(x) = 0.$$

Mnożąc pierwszą z równości przez C_1, drugą przez C_2 i dodając je stronami otrzymujemy, że dla $x \in (a,b)$

$$(C_1y_1''(x) + C_2y_2''(x)) + p_1(x)(C_1y_1'(x) + C_2y_2'(x)) + p_2(x)(C_1y_1(x) + C_2y_2(x)) = 0$$

czyli

$$(C_1y_1(x) + C_2y_2(x))'' + p_1(x)(C_1y_1(x) + C_2y_2(x))' + p_2(x)(C_1y_1(x) + C_2y_2(x)) = 0$$

a to oznacza, że funkcja $y = C_1y_1(x) + C_2y_2(x)$ jest rozwiązaniem równania (6.2).

Można sprawdzić, że $y = e^{2x}$ i $y = e^{-x}$, $x \in R$, są rozwiązaniami równania $y'' - y' - 2y = 0$. W konsekwencji każda funkcja określona wzorem

$$y = C_1e^{2x} + C_2e^{-x},$$

gdzie C_1, C_2 są dowolnymi stałymi, jest również rozwiązaniem tego równania. Powstaje pytanie: czy jest to rozwiązanie ogólne rozpatrywanego równania? W tym przykładzie rzeczywiście tak jest. W ogólności jednak, kombinacja liniowa (z dowolnymi stałymi C_1 i C_2) dwóch rozwiązań równania (6.2) n i e m u s i być rozwiązaniem ogólnym tego równania.

Aby wyjaśnić to zagadnienie, wprowadzimy pojęcie funkcji liniowo niezależnych.

Funkcje $y_1(x)$ i $y_2(x)$ nazywamy **liniowo zależnymi** na przedziale (a,b), gdy istnieją stałe C_1 i C_2, z których przynajmniej jedna jest różna od zera, takie, że

(6.3) $$\bigwedge_{x \in (a,b)} C_1y_1(x) + C_2y_2(x) = 0.$$

Funkcje $y_1(x)$ i $y_2(x)$ nazywamy **liniowo niezależnymi** na przedziale (a,b), gdy nie są liniowo zależne.

Oznacza to, że dla funkcji liniowo niezależnych warunek (6.3) jest spełniony jedynie wtedy, gdy $C_1 = 0$ i $C_2 = 0$.

Na przykład funkcje $y_1(x) = 2e^{-x}$ i $y_2(x) = 3e^{-x}$ są liniowo zależne na zbiorze R, gdyż

$$\bigwedge_{x \in R} -3y_1(x) + 2y_2(x) = 0.$$

Natomiast funkcje $y_1(x) = e^{-x}$ i $y_2(x) = e^{2x}$ są liniowo niezależne na R, gdyż przypuszczenie, że istnieją stałe C_1 i C_2, nie równe jednocześnie

zeru, dla których $C_1e^{-x}+C_2e^{2x}=0$ dla każdego $x\in R$, prowadzi do wniosku, że funkcja $y=e^{3x}$ lub $y=e^{-3x}$ jest stała, co jest nieprawdą.

Z podanej wyżej definicji liniowej niezależności łatwo wynika, że funkcje są liniowo niezależne na przedziale , gdy ich iloraz nie jest funkcją stałą na tym przedziale.

Zatem funkcje $y_1(x)=\sin x$ i $y_2(x)=e^x$ są liniowo niezależne na zbiorze R, gdyż

$$\frac{y_1(x)}{y_2(x)}=\frac{\sin x}{e^x}\quad \text{const.}$$

Dla rozwiązań równania liniowego jednorodnego (6.2) wykazuje się, że:

Rozwiązania szczególne $y_1(x)$ *i* $y_2(x)$ *równania* (6.2) *na przedziale* (a,b) *są funkcjami liniowo niezależnymi na tym przedziale wtedy i tylko wtedy, gdy*

$$W(x)=\begin{vmatrix} y_1(x) & y_2(x)\\ y_1'(x) & y_2'(x)\end{vmatrix}\neq 0$$

dla każdego $x\in(a,b)$.

Wyznacznik W(x) nazywamy wyznacznikiem Wrońskiego lub krócej: **wrońskianem** funkcji y_1 i y_2.

TWIERDZENIE 6.3. Jeżeli y_1 i y_2 są dwoma liniowo niezależnymi na przedziale (a,b) rozwiązaniami szczególnymi równania liniowego jednorodnego (6.2) oraz C_1 i C_2 są dowolnymi stałymi, to

(6.4) $$y=C_1y_1(x)+C_2y_2(x),\quad x\in(a,b),$$

jest rozwiązaniem ogólnym równania (6.2), przy czym zawiera ono wszystkie rozwiązania tego równania.

Dowód. Z twierdzenia 6.2 wynika, że funkcja postaci (6.4) jest rozwiązaniem równania (6.2) dla dowolnych C_1 i C_2. Wystarczy wykazać, że dla dowolnej trójki liczb (x_0,y_0,y_0'), gdzie $x_0\in(a,b)$, można tak dobrać stałe C_1 i C_2, by wzór (6.4) określał rozwiązanie szczególne $y=y(x)$ spełniające warunki $y(x_0)=y_0$ i $y'(x_0)=y_0'$.

Wśród rozwiązań (6.4) znajduje się takie, które spełnia warunki

$y(x_0)=y_0$ i $y'(x_0)=y_0'$ wtedy i tylko wtedy, gdy stałe C_1 i C_2 spełniają następujący układ równań:

(6.5) $$\begin{cases} C_1y_1(x_0)+C_2y_2(x_0)=y_0,\\ C_1y_1'(x_0)+C_2y_2'(x_0)=y_0'\end{cases}$$

Jest to układ liniowy względem niewiadomych C_1 i C_2, którego wyznacznik główny jest równy

$$\begin{vmatrix} y_1(x_0) & y_2(x_0)\\ y_1'(x_0) & y_2'(x_0)\end{vmatrix}=W(x_0).$$

Z liniowej niezależności rozwiązań y_1 i y_2 na przedziale (a,b) wynika, że $W(x)\neq 0$ dla $x\in(a,b)$. Zatem $W(x_0)\neq 0$. Na podstawie twierdzenia Cramera dla układów liniowych wnioskujemy więc, że układ (6.5) ma dokładnie jedno rozwiązanie i jest ono określone wzorami:

(6.6) $$C_1=\frac{1}{W(x_0)}\begin{vmatrix} y_0 & y_2(x_0)\\ y_0' & y_2'(x_0)\end{vmatrix},\qquad C_2=\frac{1}{W(x_0)}\begin{vmatrix} y_1(x_0) & y_0\\ y_1'(x_0) & y_0'\end{vmatrix}.$$

Tak więc, funkcja określona wzorem (6.4) ze stałymi (6.6) jest rozwiązaniem szczególnym równania (6.2), spełniającym warunki $y(x_0)=y_0$ i $y'(x_0)=y_0'$, co kończy dowód. □

PRZYKŁAD 6.1. Łatwo sprawdzić, że $y_1(x)=1$, $y_2(x)=e^{2x}$, $x\in R$, są rozwiązaniami szczególnymi równania

(1) $$y''-2y'=0.$$

Są to rozwiązania liniowe niezależne, gdyż wrońskian tych funkcji

$$W(x)=\begin{vmatrix} 1 & e^{2x}\\ 0 & 2e^{2x}\end{vmatrix}=2e^{2x}\neq 0\quad \text{dla}\quad x\in R.$$

Zatem zgodnie z twierdzeniem 6.3

$$y=C_1+C_2e^{2x},$$

gdzie C_1,C_2 są dowolnymi stałymi, jest rozwiązaniem ogólnym równania (1). ■

Z powyższych rozważań wynika, że dla znalezienia rozwiązania ogólnego równania liniowego jednorodnego II rzędu wystarczy znać dwa liniowo niezależne rozwiązania szczególne. Okazuje się jednak, że wyznaczenie (czy odgadnięcie) takich rozwiązań jest na ogół trudne, a czasem wręcz niemożliwe, gdyż rozwiązania te nie zawsze wyrażają się przez funkcje elementarne (por. przykład 6.2).

Jeśli znamy jedno rozwiązanie równania liniowego jednorodnego (6.2), to dla znalezienia drugiego (liniowo niezależnego) rozwiązania okazać się może pomocne następujące twierdzenie:

TWIERDZENIE 6.4. Jeżeli $y = y_1(x), x \in (a,b)$, jest rozwiązaniem szczególnym równania (6.2), przy czym $y_1(x) \neq 0$ dla $x \in (a,b)$, to funkcja określona wzorem

(6.7) $$y_2(x) = y_1(x) \int \frac{1}{y_1^2(x)} e^{-P(x)} dx,$$

gdzie $P(x)$ oznacza dowolną z funkcji pierwotnych funkcji $p_1(x)$ na przedziale (a,b), jest również rozwiązaniem równania (6.2), przy czym $y_1(x)$ i $y_2(x)$ są rozwiązaniami liniowo niezależnymi.

PRZYKŁAD 6.2.

a) Łatwo odgadnąć, że $y_1(x) = 1$, $x \in (0,+\infty)$, jest rozwiązaniem szczególnym równania

(1) $$y'' + \frac{2}{x} y' = 0, \quad x > 0.$$

Korzystając z twierdzenia 6.4 wyznaczymy drugie rozwiązanie:

$$y_2(x) = \int e^{-P(x)} dx = \left\{ P(x) \underset{C}{=} \int \frac{2}{x} dx \underset{C}{=} 2 \ln x = \ln x^2 \right\} =$$

$$= \int e^{-\ln x^2} dx = \int \frac{1}{x^2} dx \underset{C}{=} -\frac{1}{x}.$$

Rozwiązania $y_1(x) = 1$ i $y_2(x) = -\frac{1}{x}$ są liniowo niezależne na przedziale $(0,+\infty)$, zatem

$$y = C_1 - C_2 \frac{1}{x}, \quad x > 0, \quad C_1, C_2 \in R,$$

jest rozwiązaniem ogólnym równania (1).

b) Łatwo widać, że $y_1(x) = 1$, $x \in R$, jest rozwiązaniem szczególnym równania

$$y'' + 2xy' = 0.$$

W tym przypadku zastosowanie twierdzenia 6.4 nie prowadzi jednak do efektywnego wyznaczenia drugiego rozwiązania, gdyż

$$y_2(x) = \int e^{-P(x)} dx = \left\{ P(x) \underset{C}{=} \int 2x dx \underset{C}{=} x^2 \right\} = \int e^{-x^2} dx,$$

a otrzymana całka nie wyraża się przez funkcje elementarne. ■

W dalszym ciągu podamy metodę rozwiązania równania liniowego jednorodnego II rzędu w przypadku, gdy funkcje p_1 i p_2 w równaniu (6.2) są stałe.

METODA UZMIENNIANIA STAŁYCH rozwiązywania równań liniowych niejednorodnych. Weźmy pod uwagę równanie liniowe niejednorodne (6.1): $y'' + p_1(x)y' + p_2(x)y = q(x)$ i odpowiadające mu równanie jednorodne (6.2): $y'' + p_1(x)y' + p_2(x)y = 0$. Wiemy już, że rozwiązanie ogólne równania (6.2) określone jest wzorem

$$y = C_1 y_1(x) + C_2 y_2(x),$$

gdzie y_1 i y_2 są dwoma dowolnymi, liniowo niezależnymi rozwiązaniami równania (6.2) na przedziale (a,b), a C_1, C_2 są dowolnymi stałymi.

Wykazuje się, że z powyższego wzoru można uzyskać każde rozwiązanie równania (6.1) przez zastąpienie stałych C_1 i C_2 odpowiednio dobranymi funkcjami zmiennej x. Co więcej, na drodze takiego u z m i e n n i e n i a s t a ł y c h można otrzymać (analogicznie jak dla równania liniowego I rzędu) rozwiązanie ogólne równania (6.1).

Załóżmy, że $C_1(x)$ i $C_2(x)$ są takimi funkcjami, że

(6.8) $$y = C_1(x)y_1(x) + C_2(x)y_2(x), \qquad x \in (a,b),$$

jest rozwiązaniem równania (6.1). Z (6.8) otrzymujemy

$$y' = C_1'(x)y_1(x) + C_2'(x)y_2(x) + C_1(x)y_1'(x) + C_2(x)y_2'(x).$$

Ograniczymy się do znalezienia takich funkcji $C_1(x)$ i $C_2(x)$, dla których:

$$C_1'(x)y_1(x) + C_2'(x)y_2(x) = 0.$$

Wówczas

(6.8') $$y' = C_1(x)y_1'(x) + C_2(x)y_2'(x)$$

oraz

(6.8'') $$y'' = C_1'(x)y_1'(x) + C_2'(x)y_2'(x) + C_1(x)y_1''(x) + C_2(x)y_2''(x).$$

Podstawiając (6.8), (6.8') i (6.8'') do równania (6.1) i uwzględniając fakt, że y_1 i y_2 są rozwiązaniami równania (6.2), otrzymujemy

$$C_1'(x)y_1'(x) + C_2'(x)y_2'(x) = q(x).$$

Stąd wynika, że pochodne funkcji $C_1(x)$ i $C_2(x)$ spełniają układ równań

(6.9) $$\begin{cases} C_1'(x)y_1(x) + C_2'(x)y_2(x) = 0, \\ C_1'(x)y_1'(x) + C_2'(x)y_2'(x) = q(x), \end{cases}$$

który jest układem liniowym względem niewiadomych $C_1'(x)$ i $C_2'(x)$. Z liniowej niezależności rozwiązań y_1 i y_2 wynika, że wyznacznik główny tego układu

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} \neq 0$$

dla każdego $x \in (a,b)$. Zgodnie z twierdzeniem Cramera układ (6.9) ma dokładnie jedno rozwiązanie:

$$C_1'(x) = \frac{1}{W(x)}\begin{vmatrix} 0 & y_2(x) \\ q(x) & y_2'(x) \end{vmatrix}, \quad C_2'(x) = \frac{1}{W(x)}\begin{vmatrix} y_1(x) & 0 \\ y_1'(x) & q(x) \end{vmatrix}.$$

Stąd

(6.10) $$C_1(x) = \int \frac{-q(x)y_2(x)}{W(x)}dx + C_1, \quad C_2(x) = \int \frac{q(x)y_1(x)}{W(x)}dx + C_2,$$

gdzie C_1 i C_2 są dowolnymi stałymi.

Wykazuje się, że wzór (6.8) z funkcjami $C_1(x)$ i $C_2(x)$ określonymi wzorami (6.10) określa rozwiązanie ogólne równania (6.1).

PRZYKŁAD 6.3. Weźmy pod uwagę równanie

(1) $$y'' + \frac{4}{x}y' + \frac{2}{x^2}y = \frac{e^x}{x^2} \quad \text{dla} \quad x > 0.$$

Można sprawdzić, że funkcje $y = \frac{1}{x}$, $x > 0$ i $y = \frac{1}{x^2}$, $x > 0$, są rozwiązaniami szczególnymi równania jednorodnego

(2) $$y'' + \frac{4}{x}y' + \frac{2}{x^2}y = 0, \quad x > 0,$$

przy czym są to rozwiązania liniowo niezależne (gdyż ich iloraz nie jest stały). Zgodnie z twierdzeniem 6.3 wzór

$$y_0 = C_1\frac{1}{x} + C_2\frac{1}{x^2}, \quad x > 0, \ C_1, C_2 \in R,$$

określa rozwiązanie ogólne równania (2).

Rozwiązanie ogólne równania (1) znajdziemy w postaci

$$y = C_1(x)\frac{1}{x} + C_2(x)\frac{1}{x^2},$$

przy czym pochodne niewiadomych funkcji $C_1(x)$ i $C_2(x)$ wyznaczymy z układu równań (por. (6.9)):

$$\begin{cases} C_1'(x)\frac{1}{x} + C_2'(x)\frac{1}{x^2} = 0, \\ -C_1'(x)\frac{1}{x^2} - C_2'(x)\frac{2}{x^3} = \frac{e^x}{x^2}. \end{cases}$$

Z układu tego otrzymujemy, że

$$C_1'(x) = e^x, \quad C_2'(x) = -xe^x,$$

a stąd

$$C_1(x) = e^x + C_1, \quad C_2(x) = e^x - xe^x + C_2,$$

gdzie C_1 i C_2 oznaczają dowolne stałe. Zatem rozwiązanie ogólne równania (1) określone jest wzorem

(3) $$y = (e^x + C_1)\frac{1}{x} + (e^x - xe^x + C_2)\frac{1}{x^2}, \quad x > 0, \ C_1, C_2 \in R,$$

Z twierdzenia 6.1 wynika, że przez każdy punkt obszaru $D = \{(x,y) \in R^2 : x > 0\}$ przechodzi dokładnie jedno rozwiązanie o zadanej wartości pochodnej. Wyznaczymy, na przykład, to rozwiązanie równania (1), dla którego $y(1) = e - 1$ oraz $y'(1) = 2 - e$. Z wzoru (3) mamy

(4) $$y' = \frac{1}{x^3}(xe^x - 2e^x - C_1x - 2C_2).$$

Podstawiając $x = 1$, $y = e - 1$, $y' = 2 - e$ we wzorach (3) i (4) otrzymujemy

$$C_1 + C_2 = -1, \quad -C_1 - 2C_2 = 2,$$

a stąd $C_1 = 0$, $C_2 = -1$. Zatem rozwiązaniem szczególnym równania (1) spełniającym podane warunki początkowe jest funkcja

$$y=\frac{e^x-1}{x^2}, \qquad x>0.$$ ■

Dla równania liniowego II rzędu, podobnie jak dla równania liniowego I rzędu, prawdziwe jest następujące twierdzenie.

TWIERDZENIE 6.5. Jeżeli $y=y_0(x,C_1,C_2)$, $C_1,C_2\in R$, jest rozwiązaniem ogólnym równania liniowego jednorodnego (6.2) na przedziale (a,b) oraz $y=y_s(x)$, $x\in(a,b)$, jest rozwiązaniem szczególnym równania niejednorodnego (6.1), to

$$y=y_0(x,C_1,C_2)+y_s(x)$$

jest rozwiązaniem ogólnym równania (6.1).

Analogicznie jak dla równania liniowego I rzędu, rozwiązanie ogólne równania niejednorodnego można wyznaczyć na podstawie twierdzenia 6.5, wykorzystując metodę uzmienniania stałych jedynie do znalezienia rozwiązania szczególnego równania niejednorodnego.

PRZYKŁAD 6.4. Rozwiążemy równanie

(1) $$y''+\frac{2}{x}y'=12x,$$

a następnie wyznaczymy rozwiązanie takie, że $y(-1)=3$, $y(-2)=-4$.

Z twierdzenia 6.1 wynika, że przez każdy punkt półpłaszczyzn

$$D_1=\{(x,y)\in R^2: x>0\}, \quad D_2=\{(x,y)\in R^2: x<0\}$$

przechodzi dokładnie jedno rozwiązanie. Szukane rozwiązanie szczególne przebiega w obszarze D_2.

Można sprawdzić, że funkcje $y=1$ i $y=\frac{1}{x}$ dla $x\neq 0$ są rozwiązaniami szczególnymi równania jednorodnego

(2) $$y''+\frac{2}{x}y'=0,$$

przy czym są to rozwiązania liniowo niezależne. Zatem wzór

$$y_0=C_1+C_2\frac{1}{x}, \qquad C_1,C_2\in R,$$

określa rozwiązanie ogólne równania (2).

Następnie metodą uzmienniania stałych znajdujemy rozwiązanie szczególne równania (1). Szukamy tego rozwiązania w postaci

$$y_s=C_1(x)+C_2(x)\frac{1}{x}.$$

Z układu

$$\begin{cases} C_1'(x)+C_2'(x)\frac{1}{x}=0, \\ -C_2'(x)\frac{1}{x^2}=12x. \end{cases}$$

wyznaczamy: $C_1'(x)$, $C_2'(x)$, a następnie przez całkowanie otrzymujemy

$$C_1(x)\underset{C}{=}4x^3, \quad C_2(x)\underset{C}{=}-3x^4.$$

Zatem

$$y_s=x^3 \quad \text{dla } x>0 \text{ lub } x<0$$

jest rozwiązaniem szczególnym równania (1).

Z twierdzenia 6.5 wynika, że $y=y_0+y_s$, czyli

(3) $$y=C_1+C_2\frac{1}{x}+x^3, \qquad C_1,C_2\in R$$

jest rozwiązaniem ogólnym równania (1).

Wyznaczymy teraz rozwiązanie szczególne spełniające warunki brzegowe $y(-1)=3$, $y(-2)=-4$. Uwzględniając te warunki w (3) otrzymujemy

$$3=C_1-C_2-1, \qquad -4=C_1-\frac{1}{2}C_2-8,$$

a stąd $C_1=4$ i $C_2=0$. Zatem szukanym rozwiązaniem jest funkcja

$$y=4+x^3, \qquad x<0.$$ ■

RÓWNANIE LINIOWE JEDNORODNE O STAŁYCH WSPÓŁCZYNNIKACH. Załóżmy, że w równaniu liniowym jednorodnym (6.2): $y''+p_1(x)y'+p_2(x)y=0$ funkcje p_1 i p_2 są stałe. Załóżmy, że $p_1(x)=a_1=\text{const.}$ i $p_2(x)=a_2=\text{const.}$ Równanie (6.2) przyjmuje postać

(6.11) $$y''+a_1y'+a_2y=0$$

i nazywane jest **równaniem liniowym jednorodnym II rzędu o stałych współczynnikach**. Całkowanie takiego równania można zawsze przeprowadzić za pomocą funkcji elementarnych, co nie jest zawsze możliwe przy dowolnych funkcjach $p_1(x)$ i $p_2(x)$ (por. przykład 6.2). Ponadto rozwiązanie równania (6.11) nie wymaga całkowania, a sprowadza się

jedynie do pewnych operacji algebraicznych. Zauważmy także, że każde rozwiązanie równania (6.11) jest określone na całym zbiorze R, czyli dla $x \in (-\infty, +\infty)$.

Wiadomo z wcześniejszych rozważań, że znając dwa liniowo niezależne rozwiązania szczególne tego równania możemy natychmiast napisać jego rozwiązanie ogólne.

Z uwagi na postać równania (6.11), rozwiązań szczególnych tego równania szukamy wśród funkcji wykładniczych postaci

$$y = e^{rx}, \qquad x \in R.$$

Po obliczeniu y', y'' i wstawieniu do (6.11) okazuje się, że funkcja $y = e^{rx}$ jest rozwiązaniem tego równania wtedy i tylko wtedy, gdy współczynnik r spełnia tzw. **równanie charakterystyczne**:

(6.12) $$r^2 + a_1 r + a_2 = 0.$$

Rozważymy kolejno trzy przypadki: $\Delta = a_1^2 - 4a_2 > 0$, $\Delta = 0$, $\Delta < 0$.

A. Jeżeli $\Delta > 0$, równanie charakterystyczne (6.12) ma dwa różne pierwiastki rzeczywiste r_1 i r_2. Zatem funkcje

$$y_1(x) = e^{r_1 x}, \qquad y_2(x) = e^{r_2 x}$$

są rozwiązaniami szczególnymi równania (6.11), przy czym są to rozwiązania liniowo niezależne, gdyż

$$W(x) = \begin{vmatrix} e^{r_1 x} & e^{r_2 x} \\ r_1 e^{r_1 x} & r_2 e^{r_2 x} \end{vmatrix} = (r_2 - r_1) e^{(r_1 + r_2)x} \neq 0$$

dla każdego $x \in R$. Zgodnie z twierdzeniem 6.3

$$y = C_1 e^{r_1 x} + C_2 e^{r_2 x}, \qquad x \in R, \ C_1, C_2 \in R,$$

jest rozwiązaniem ogólnym równania (6.11).

PRZYKŁAD 6.5. Znajdziemy rozwiązanie ogólne równania

(1) $$y'' + 2y' - 3y = 0.$$

Ponieważ równanie charakterystyczne

$$r^2 + 2r - 3 = 0$$

ma dwa pierwiastki: $r_1 = 1$ i $r_2 = -3$, więc

$$y = C_1 e^{x} + C_2 e^{-3x}, \qquad x \in R, \ C_1, C_2 \in R,$$

jest rozwiązaniem ogólnym równania (1). ■

B. Jeżeli $\Delta = 0$, to równanie charakterystyczne (6.12) ma jeden (podwójny) pierwiastek: $r_0 = -\frac{1}{2}a_1$. Funkcja $y_1(x) = e^{r_0 x}$ jest rozwiązaniem szczególnym równania (6.11). Wykażemy, że w tym przypadku funkcja $y_2(x) = xe^{r_0 x}$ jest również rozwiązaniem tego równania. Obliczamy

$$y_2'(x) = (1 + r_0 x)e^{r_0 x}, \qquad y_2''(x) = r_0(2 + r_0 x)e^{r_0 x},$$

wstawiamy do równania (6.11) i otrzymujemy:

$$e^{r_0 x}((r_0^2 + r_0 a_1 + a_2)x + 2r_0 + a_1) = 0.$$

Ponieważ $r_0^2 + r_0 a_1 + a_2 = 0$ (r_0 jest pierwiastkiem równania (6.12)) i $r_0 = -\frac{1}{2}a_1$, więc równanie jest spełnione dla każdego $x \in R$. Oznacza to, że funkcja $y_2(x) = xe^{r_0 x}$ jest rozwiązaniem równania (6.11).

Rozwiązania $y_1(x)$ i $y_2(x)$ są liniowo niezależne, gdyż

$$W(x) = \begin{vmatrix} e^{r_0 x} & xe^{r_0 x} \\ r_0 e^{r_0 x} & (1 + r_0 x)e^{r_0 x} \end{vmatrix} = e^{2r_0 x} \neq 0$$

dla każdego $x \in R$.

Zgodnie z twierdzeniem 6.3

$$y = C_1 e^{r_0 x} + C_2 x e^{r_0 x}, \qquad x \in R, \ C_1, C_2 \in R,$$

jest rozwiązaniem ogólnym równania (6.11).

PRZYKŁAD 6.6. Znajdziemy rozwiązanie ogólne równania

(1) $$y'' + 2y' + y = 0.$$

Ponieważ równanie charakterystyczne

$$r^2 + 2r + 1 = 0$$

ma pierwiastek podwójny $r_0 = -1$, więc rozwiązanie ogólne równania (1) określa wzór

$$y = C_1 e^{-x} + C_2 x e^{-x}, \qquad x \in R, \ C_1, C_2 \in R. \quad ■$$

C. Jeżeli $\Delta < 0$, to równanie charakterystyczne (6.12) ma pierwiastki zespolone sprzężone: $r_1 = \alpha + i\beta$, $r_2 = \alpha - i\beta$, gdzie $\alpha = -\frac{1}{2}a_1$, $\beta = \frac{1}{2}\sqrt{-\Delta} \neq 0$. Okazuje się wówczas, że rozwiązaniami szczególnymi równania (6.11)są funkcje

$$y_1(x) = e^{\alpha x}\cos\beta x, \qquad y_2(x) = e^{\alpha x}\sin\beta x.$$

Obliczamy

$$y_1'(x) = e^{\alpha x}(\alpha\cos\beta x - \beta\sin\beta x),$$

$$y_1''(x) = e^{\alpha x}((\alpha^2 - \beta^2)\cos\beta x - 2\alpha\beta\sin\beta x)$$

i podstawiamy do równania (6.11). Po uporządkowaniu mamy

$$(\alpha^2 - \beta^2 + a_1\alpha + a_2)\cos\beta x - (2\alpha\beta + a_1\beta)\sin\beta x = 0.$$

Uwzględniając, że $\alpha = -\frac{1}{2}a_1$, $\beta = \frac{1}{2}\sqrt{4a_2 - a_1^2}$ stwierdzamy łatwo, że równość zachodzi dla każdego $x \in R$. Analogicznie wykazuje się, że y_2 jest rozwiązaniem równania (6.11).

Rozwiązania $y_1(x)$ i $y_2(x)$ są liniowo niezależne, gdyż

$$W(x) = \begin{vmatrix} e^{\alpha x}\cos\beta x & e^{\alpha x}\sin\beta x \\ e^{\alpha x}(\alpha\cos\beta x - \beta\sin\beta x) & e^{\alpha x}(\alpha\sin\beta x + \beta\cos\beta x) \end{vmatrix} = \beta e^{\alpha x} \neq 0$$

dla każdego $x \in R$.

Zgodnie z twierdzeniem 6.3

$$y = e^{\alpha x}(C_1\cos\beta x + C_2\sin\beta x), \qquad x \in R,\ C_1, C_2 \in R,$$

jest rozwiązaniem ogólnym równania (6.11).

U w a g a . Jeżeli równanie charakterystyczne ma pierwiastki zespolone $r = \alpha \pm i\beta$, to funkcje zespolone

$$Y_1(x) = e^{(\alpha+i\beta)x} = e^{\alpha x}(\cos\beta x + i\sin\beta x),$$

$$Y_2(x) = e^{(\alpha-i\beta)x} = e^{\alpha x}(\cos\beta x - i\sin\beta x)$$

są rozwiązaniami równania (6.11). Ponieważ jednak rozważamy tu równania różniczkowe o współczynnikach rzeczywistych, oczekujemy rozwiązań także rzeczywistych. W ogólnej teorii równań liniowych łatwo wykazuje się, że:

jeżeli funkcja zespolona zmiennej rzeczywistej

$$Y(x) = U(x) + iV(x),$$

gdzie $U(x) = \mathrm{Re}\,Y(x)$, $V(x) = \mathrm{Im}\,Y(x)$, jest rozwiązaniem równania liniowego jednorodnego o współczynnikach rzeczywistych, to funkcje $U(x)$ oraz $V(x)$ są również rozwiązaniami tego równania.

Korzystając z tego twierdzenia otrzymujemy w tym przypadku, że $y_1(x) = e^{\alpha x}\cos\beta x$ oraz $y_2(x) = e^{\alpha x}\sin\beta x$ są rozwiązaniami równania (6.11)

P R Z Y K Ł A D 6.7. Znajdziemy rozwiązania ogólne równań:

a) $y'' + 4y = 0$, b) $y'' - 2y' + 10y = 0$.

a) Równanie charakterystyczne $r^2 + 4 = 0$ ma pierwiastki zespolone $r_1 = 2i$, $r_2 = -2i$ ($\alpha = 0$, $\beta = 2$). Rozwiązanie ogólne danego równania różniczkowego określone jest wzorem

$$y = C_1\cos 2x + C_2\sin 2x, \qquad x \in R,\ C_1, C_2 \in R.$$

b) Równanie charakterystyczne

$$r^2 - 2r + 10 = 0$$

ma pierwiastki zespolone $r_1 = 1 + 3i$, $r_2 = 1 - 3i$ ($\alpha = 1$, $\beta = 3$). Rozwiązanie ogólne rozważanego równania różniczkowego ma postać

$$y = e^x(C_1\cos 3x + C_2\sin 3x), \qquad x \in R,\ C_1, C_2 \in R. \quad \blacksquare$$

Wyniki naszych rozważań w punktach A, B i C można zapisać w tabeli:

Pierw. równ. charakter. (6.12)	Rozw. ogólne równania (6.11)
$\Delta > 0$, r_1, r_2	$y = C_1e^{r_1x} + C_2e^{r_2x}$
$\Delta = 0$, r_0	$y = C_1e^{r_0x} + C_2xe^{r_0x}$
$\Delta < 0$, $r = \alpha \pm i\beta$	$y = e^{\alpha x}(C_1\cos\beta x + C_2\sin\beta x)$

RÓWNANIE LINIOWE NIEJEDNORODNE O STAŁYCH WSPÓŁCZYNNIKACH. Równanie liniowe (6.1), w którym $p_1(x) = a_1 = \mathrm{const}$, $p_2(x) = a_2 = \mathrm{const}$, przyjmuje postać

$$y'' + a_1y' + a_2y = q(x) \tag{6.13}$$

i jest nazywane **równaniem liniowym II rzędu o stałych współczynnikach**.

Ponieważ umiemy już dla dowolnych $a_1, a_2 \in R$ znaleźć rozwiązanie ogólne równania jednorodnego $y'' + a_1 y' + a_2 y = 0$, więc stosując poznaną już metodę uzmienniania stałych możemy znaleźć rozwiązanie ogólne równania liniowego niejednorodnego (6.13).

PRZYKŁAD 6.8. Rozwiążemy równanie

(1) $$y'' + y' = \frac{1}{1+e^x}.$$

Najpierw rozwiążemy równanie jednorodne

(2) $$y'' + y' = 0.$$

Ponieważ równanie charakterystyczne

$$r^2 + r = 0$$

ma dwa pierwiastki: $r = 0$, $r = -1$, więc rozwiązanie ogólne równania jednorodnego (2) określone jest wzorem:

$$y_0 = C_1 + C_2 e^{-x}, \quad x \in R, \quad C_1, C_2 \in R.$$

Szukamy rozwiązania szczególnego równania (1) w postaci

$$y_s = C_1(x) + C_2(x)e^{-x}.$$

Niewiadome funkcje $C_1(x)$ i $C_2(x)$ wyznaczymy z układu równań (por. (6.9))

$$\begin{cases} C_1'(x) + C_2'(x)e^{-x} = 0, \\ -C_2'(x)e^{-x} = \dfrac{1}{1+e^x}. \end{cases}$$

Kolejno otrzymujemy

$$C_1'(x) = \frac{1}{1+e^x}, \qquad C_2'(x) = -\frac{e^x}{1+e^x},$$

$$C_1(x) = \int \frac{1}{1+e^x}dx = \int (1 - \frac{e^x}{1+e^x})dx \underset{C}{=} x - \ln(1+e^x),$$

$$C_2(x) = \int -\frac{e^x}{1+e^x}dx \underset{C}{=} -\ln(1+e^x),$$

a następnie

$$y_s = x - \ln(1+e^x) - e^{-x}\ln(1+e^x),$$

czyli

$$y_s = x - (1+e^{-x})\ln(1+e^x).$$

Zgodnie z twierdzeniem 6.5

$$y = y_0 + y_s = C_1 + C_2 e^{-x} + x - (1+e^{-x})\ln(1+e^x), \quad x \in R,$$

gdzie C_1, C_2 są dowolnymi stałymi, jest rozwiązaniem ogólnym równania (1). ■

METODA PRZEWIDYWANIA dla równań liniowych niejednorodnych o stałych współczynnikach. Weźmy pod uwagę równanie (6.13): $y'' + a_1 y' + a_2 y = q(x)$. W pewnych przypadkach, podobnie jak dla równania liniowego I rzędu, potrafimy przewidzieć postać rozwiązania szczególnego tego równania i wyznaczyć takie rozwiązanie z pominięciem całkowania (które zawsze występuje przy stosowaniu metody uzmienniania stałych).

I. *Jeżeli w równaniu* (6.13)

$$q(x) = W_n(x)e^{\alpha x},$$

gdzie $\alpha \in R$, *a* W_n *jest wielomianem stopnia* n, *to istnieje rozwiązanie szczególne równania* (6.13) *postaci*

$$y_s = x^k P_n(x) e^{\alpha x},$$

gdzie P_n *jest pewnym wielomianem stopnia* n, *zaś* k - *krotnością pierwiastka* α *równania charakterystycznego (gdy* α *nie jest pierwiastkiem równania charakterystycznego, przyjmujemy* $k = 0$*).*

Na przykład:

a) dla równania $y' - y = x^2 + 1$ istnieje rozwiązanie szczególne postaci $y_s = Ax^2 + Bx + C$

b) dla równania $y'' - y = xe^{2x}$ przewidujemy, że $y_s = (Ax + B)e^{2x}$;

c) dla równania $y'' - y = (x+1)e^{-x}$ przewidujemy rozwiązanie szczególne w postaci $y_s = x(Ax + B)e^{-x}$, gdyż $\alpha = -1$ jest pierwiastkiem jednokrotnym równania charakterystycznego $r^2 - 1 = 0$;

d) dla równania $y''-2y'+y=3e^x$ istnieje rozwiązanie szczególne postaci $y_s=x^2Ae^x$, gdyż $\alpha=1$ jest dwukrotnym pierwiastkiem równania charakterystycznego $r^2-2r+1=0$.

Niewiadome współczynniki A,B i C w tych przykładach wyznaczamy korzystając z faktu, że y_s spełnia odpowiednie równanie różniczkowe.

II. *Jeżeli w równaniu* (6.13)

$$q(x)=W_n(x)\cos\beta x+V_m(x)\sin\beta x,$$

przy czym $\beta\in R$, *a* W_n, V_m *są wielomianami odpowiednich stopni, to dla równania* (6.13) *istnieje rozwiązanie szczególne postaci*

$$y_s=x^k(P_l(x)\cos\beta x+Q_l(x)\sin\beta x),$$

gdzie P_l, Q_l *są pewnymi wielomianami stopnia* l, $l=\max(n,m)$, *oraz* $k=1$, *gdy* $r=\beta i$ *jest pierwiastkiem lub* $k=0$, *gdy* $r=\beta i$ *nie jest pierwiastkiem równania charakterystycznego* (6.12).

Na przykład:

a) dla równania $y''-y=x\sin3x+\cos3x$ istnieje rozwiązanie szczególne postaci $y_s=(Ax+B)\sin3x+(Cx+D)\cos3x$;

b) dla równania $y''-y=5\sin2x$ przewidujemy rozwiązanie szczególne postaci $y_s=A\sin2x+B\cos2x$;

c) dla równania $y''+4y=5\sin2x$ istnieje rozwiązanie szczególne postaci $y_s=x(A\sin2x+B\cos2x)$, gdyż (w odróżnieniu od przykładu b)) $r=\beta i=2i$ jest pierwiastkiem równania charakterystycznego.

III. *Jeżeli w równaniu* (6.13)

$$q(x)=(W_n(x)\cos\beta x+V_m(x)\sin\beta x)e^{\alpha x},$$

przy czym $\alpha\in R$, $\beta\in R$, W_n, V_m *są wielomianami odpowiednich stopni, to istnieje rozwiązanie szczególne równania* (6.13) *postaci*

$$y_s=x^k(P_l(x)\cos\beta x+Q_l(x)\sin\beta x)e^{\alpha x},$$

gdzie P_l, Q_l *są pewnymi wielomianami stopnia* l, $l=\max(n,m)$, *zaś* k *oznacza krotność pierwiastka* $r=\alpha+i\beta$ *równania charakterystycznego* (6.12).

Zauważmy, że z tego najogólniejszego przypadku otrzymujemy w szczególności dwa wcześniej rozważane: I, gdy $\beta=0$ lub II, gdy $\alpha=0$.

Przy przewidywaniu rozwiązania szczególnego równania (6.13) korzystać możemy również z następującego twierdzenia:

TWIERDZENIE 6.6. Jeżeli $y=y_1(x)$, $x\in(a,b)$, jest rozwiązaniem równania

$$y''+a_1y'+a_2y=q_1(x),$$

a $y=y_2(x)$, $x\in(a,b)$, jest rozwiązaniem równania

$$y''+a_1y'+a_2y=q_2(x),$$

to $y=y_1(x)+y_2(x)$ jest rozwiązaniem równania

$$y''+a_1y'+a_2y=q_1(x)+q_2(x).$$

Z twierdzenia tego wynika, że dla równań

a) $y''-y'=x-e^{2x}$, b) $y''+y=xe^{-x}+2\cos x$

istnieją rozwiązania szczególne odpowiednio postaci:

a) $y_s=x(Ax+B)+Ce^{2x}$,

b) $y_s=(Ax+B)e^{-x}+x(C\cos x+D\sin x)$,

gdzie A, B, C, D są odpowiednio dobranymi współczynnikami.

PRZYKŁAD 6.9. Znajdziemy rozwiązanie ogólne równania

$$y''-2y'=2x-6x^2, \tag{1}$$

a następnie rozwiązanie szczególne spełniające warunki początkowe $y(0)=0$, $y'(0)=-1$.

Rozwiązujemy najpierw równanie jednorodne

$$y''-2y'=0. \tag{2}$$

Ponieważ równanie charakterystyczne

$$r^2-2r=0$$

ma dwa pierwiastki $r_1=0$, $r_2=2$, więc rozwiązanie ogólne równania (2) jest określone wzorem:

$$y_0=C_1+C_2e^{2x},\quad x\in R,\ C_1,C_2\in R.$$

Rozwiązanie szczególne równania (1) przewidujemy w postaci

$$y_s=x(Ax^2+Bx+C)$$

($\alpha=0$ jest pierwiastkiem równania charakterystycznego).

Po podstawieniu y_s, y_s', y_s'' do równania (1) otrzymujemy

$$6Ax+2B-6Ax^2-4Bx-2C\equiv 2x-6x^2,$$

skąd wynika, że $A=1$, $B=1$, $C=1$. Zatem

$$y_s=x^3+x^2+x, \quad x\in R,$$

jest rozwiązaniem szczególnym równania (1), a

$$(3)\quad y=y_0+y_s=C_1+C_2e^{2x}+x^3+x^2+x, \quad x\in R,\ C_1,C_2\in R,$$

jest rozwiązaniem ogólnym tego równania.

Wyznaczymy teraz rozwiązanie szczególne spełniające warunki $y(0)=0$, $y'(0)=-1$. Z (3) otrzymamy

$$(4)\qquad y'=2C_2e^{2x}+3x^2+2x+1.$$

Uwzględniając warunki początkowe w (3) i (4) mamy

$$y(0)=C_1+C_2=0, \qquad y'(0)=2C_2+1=-1,$$

a stąd $C_1=1$, $C_2=-1$. Szukane rozwiązanie szczególne określone jest wzorem

$$y=1-e^{2x}+x^3+x^2+x, \qquad x\in R. \quad \blacksquare$$

PRZYKŁAD 6.10. Rozwiążemy równanie

$$(1)\qquad y''+y=x^2-4\sin x.$$

Rozwiązujemy równanie jednorodne

$$(2)\qquad y''+y=0.$$

Ponieważ równanie charakterystyczne

$$r^2+1=0$$

ma dwa pierwiastki zespolone $r=\pm i$, więc rozwiązanie ogólne równania (2) określone jest wzorem

$$y_0=C_1\cos x+C_2\sin x, \quad x\in R,\ C_1,C_2\in R.$$

Rozwiązanie szczególne równania (1) wyznaczymy jako sumę rozwiązań szczególnych równań

$$(3)\quad y''+y=x^2 \qquad \text{i} \qquad (4)\quad y''+y=-4\sin x.$$

Rozwiązanie szczególne równania (3) przewidujemy w postaci

$$y_1(x)=Ax^2+Bx+C.$$

Obliczamy: $y_1'=2Ax+B$, $y_1''=2A$, podstawiamy do (3) i otrzymujemy

$$2A+Ax^2+Bx+C\equiv x^2.$$

Stąd wynika, że $A=1$, $B=0$, $C=-2$. Tak więc, $y_1(x)=x^2-2$. Rozwiązanie równania (4) przewidujemy w postaci

$$y_2(x)=x(A\cos x+B\sin x)$$

($r=i$ jest pierwiastkiem równania charakterystycznego). Wtedy

$y_2'=(A+Bx)\cos x+(B-Ax)\sin x$, $y_2''=(2B-Ax)\cos x+(-2A-Bx)\sin x$.

Po podstawieniu do (4) i uporządkowaniu otrzymujemy

$$2B\cos x-2A\sin x\equiv -4\sin x,$$

skąd $A=2$, $B=0$. Zatem $y_2(x)=2x\cos x$. Z twierdzenia 6.6 wynika, że

$$y_s=y_1(x)+y_2(x)=x^2-2+2x\cos x, \qquad x\in R,$$

jest rozwiązaniem szczególnym równania (1).

Rozwiązanie ogólne równania (1) określa wzór $y=y_0+y_s$, czyli

$$y=C_1\cos x+C_2\sin x+x^2-2+2x\cos x, \quad x\in R,\ C_1,C_2\in R. \quad \blacksquare$$

ZADANIA DO ROZWIĄZANIA

1. Znaleźć rozwiązanie ogólne równania wiedząc, że $y=y_1(x)$ jest jego rozwiązaniem szczególnym:

   a) $y''+\frac{1}{x}y'-\frac{1}{x^2}y=0$, $x>0$; $y_1(x)=\frac{1}{x}$,

   b) $y''-\frac{2x}{1+x^2}y'+\frac{2}{1+x^2}y=0$; $y_1(x)=x$,

   c) $y''+\frac{1}{x}y'-\frac{4}{x^2}y=0$, $x>0$; $y_1(x)=x^2$.

2. Znaleźć rozwiązanie ogólne równania

$$y''-\frac{1}{x}y'+\frac{1}{x^2}y=\frac{2}{x}, \quad x>0,$$

wiedząc, że $y_1(x) = x$ jest rozwiązaniem odpowiadającego mu równania jednorodnego. Znaleźć rozwiązanie szczególne spełniające warunki początkowe $y(1) = 0$, $y'(1) = 1$.

3. Wiedząc, że $y_1(x) = x^2$, $x > 0$, jest rozwiązaniem równania $y'' + \frac{1}{x}y' - \frac{4}{x^2}y = 0$, znaleźć rozwiązanie ogólne równania

$$y'' + \frac{1}{x}y' - \frac{4}{x^2}y = \frac{12}{x^3}, \quad x > 0,$$

oraz rozwiązanie szczególne, dla którego $y(2) = -1$, $y'(2) = 0$.

4. Rozwiązać równanie:
a) $y'' + 2y' - 3y = 0$, b) $y'' - y = 0$, c) $y'' + 2y' = 0$,
d) $y'' - 2y' + y = 0$, e) $y'' + y = 0$, f) $y'' + 4y = 0$,
g) $y'' - 4y' + 4y = 0$, h) $y'' - 2y' + 10y = 0$, i) $y'' - 4y' = 0$.

5. Rozwiązać równanie $y'' - y' - 2y = 0$, a następnie
a) znaleźć wszystkie rozwiązania $y = y(x)$ równania, dla których $y(0) = 0$,
b) podać trzy funkcje $y = y(x)$, które są rozwiązaniami tego równania i spełniają warunek $y(0) = 1$,
c) znaleźć rozwiązanie spełniające warunki początkowe $y(0) = 0$, $y'(0) = -3$.

6. Rozwiązać równanie $y'' + 2y' = 0$, a następnie
a) znaleźć wszystkie rozwiązania $y = y(x)$ tego równania, dla których $\lim_{x \to +\infty} y(x) = 1$,
b) znaleźć wszystkie rozwiązania $y = y(x)$ tego równania przechodzące przez punkt $(0,0)$, dla których $\lim_{x \to -\infty} y(x) = +\infty$.

7. Stosując metodę uzmienniania stałych rozwiązać równanie:
a) $y'' - y' = 2x$, b) $y'' + y' - 6y = 5e^x$,
c) $y'' + y = 4\sin x$, d) $y'' - 2y' + y = e^{2x} - 1$,
e) $y'' - y = \frac{2e^x}{1+e^x}$, f) $y'' - 2y' + y = \frac{e^x}{4+x^2}$.

8. Stosując metodę przewidywania znaleźć rozwiązanie szczególne równania:
a) $y'' - y = 2e^{-x}$, b) $y'' + y' - 2y = 9 - 4x^3$,
c) $y'' + y = 4xe^x$, d) $y'' - y' - 2y = 3xe^{2x}$,
e) $y'' - y' = 3$, f) $y'' - y' = -2x\cos x - 2\sin x$,
g) $y'' + y' = 2e^{-x}$, h) $y'' - 2y' + 2y = 2x - e^x$,
i) $y'' + y = 2\cos x$, j) $y'' + 2y' + y = 2e^{-x} + 6\sin x$.

9. Stosując metodę przewidywania rozwiązać równanie:
a) $y'' + 2y' - 3y = 3x - 2$, b) $y'' + 4y' + 4y = 5\cos x$,
c) $y'' - 2y' + y = (3x+2)e^x$, d) $y'' + 2y' = 6e^{-2x}$,
e) $y'' - 2y' + 2y = 4x - 2x^2$, f) $y'' - y' = e^{2x} - 4x$.

10. Rozwiązać równanie:
a) $y'' + y' = 2\cos x$, b) $y'' - 2y' + 10y = 5x + 4$,
c) $y'' + y = \frac{1}{\sin^3 x}$, d) $y'' - 2y' + y = \frac{2e^x}{x^3}$,
e) $y'' - y = \frac{4x}{e^x}$, f) $y'' + 2y' + y = \frac{2\ln x}{xe^x}$,
g) $y'' - y' = \frac{1 - e^{2x}}{1 + e^x}$, h) $y'' + y = \frac{\cos x}{\sin^2 x}$,
i) $y'' - 4y = 3e^x$, j) $y'' + y = \frac{1}{\sin x}$,
k) $y'' + 2y' = 3x^2 + 3x$, l) $y'' - 3y' + 2y = 8\cos^2 x$.

11. Rozwiązać równanie przy podanych warunkach:
a) $y'' + 2y' + y = 2e^{-x}$, $y(0) = 1$, $y'(0) = 1$,
b) $y'' + y = x - 3x^2$, $y(0) = 3$, $y'(0) = 1$,
c) $y'' + 2y' + y = \frac{2e^{-x}}{1+x^2}$, $y(0) = 2$, $y'(0) = -2$,
d) $y'' + y = \frac{1}{\cos x}$, $y(\pi) = -1$, $y'(\pi) = -\pi$,

c) $y''+y=3\sin 2x$, $y(0)=1$, $y(\pi/2)=0$,

f) $y''-2y'+y=(x+1)e^{2x}$, $y(0)=-1$, $y(1)=2e$.

12. Rozwiązać równanie $y''-2y'+y=x^2-6$, a następnie znaleźć wszystkie rozwiązania tego równania, dla których $y(0)=1$.

13. Sprawdzić, że dla każdego rozwiązania $y=y(x)$ równania $y''+2y'+y=3e^{-2x}$ zachodzi warunek $\lim\limits_{x\to+\infty} y(x)=0$.

14. Rozwiązać równanie $y''-y=8e^{-3x}$, a następnie wyznaczyć wszystkie rozwiązania $y=y(x)$ tego równania, dla których $y(0)=1$ i $\lim\limits_{x\to+\infty} y(x)=-\infty$.

15. Rozwiązać równanie $y''-2y'+y=2e^x$, a następnie
a) znaleźć rozwiązanie spełniające warunki $y(1)=-3e$, $y'(1)=0$, naszkicować wykres tego rozwiązania,
b) znaleźć wszystkie rozwiązania tego równania, dla których $y(0)=1$ oraz $y(x)\geq 0$ dla każdego $x\in R$; naszkicować wykresy dwóch takich rozwiązań.

Odpowiedzi. Litery C_1 i C_2 występujące w odpowiedziach oznaczają stałe, które przyjmują, jeśli nie ma innej informacji, dowolne wartości rzeczywiste.

1. a) $y=C_1\frac{1}{x}+C_2x$, $x>0$, b) $y=C_1x+C_2(x^2-1)$, c) $y=C_1x^2+C_2\frac{1}{x^2}$, $x>0$.
2. $y=C_1x+C_2x\ln x+x\ln^2 x$, $y_s=x\ln x(1+\ln x)$.
3. $y=C_1x^2+C_2\frac{1}{x^2}-\frac{4}{x}$, $y_s=\frac{4}{x^2}-\frac{4}{x}$, $x>0$.
4. a) $y=C_1e^x+C_2e^{-3x}$, b) $y=C_1e^x+C_2e^{-x}$, c) $y=C_1+C_2e^{-2x}$,
d) $y=C_1e^x+C_2xe^x$, e) $y=C_1\cos x+C_2\sin x$, f) $y=C_1\cos 2x+C_2\sin 2x$,
g) $y=C_1e^{2x}+C_2xe^{2x}$, h) $y=e^x(C_1\cos 3x+C_2\sin 3x)$, i) $y=C_1+C_2e^{2x}$.
5. $y=C_1e^{-x}+C_2e^{2x}$, a) $y=C_1e^{-x}-C_1e^{2x}$, b) $y=C_1e^{-x}+(1-C_1)e^{2x}$,
np.: $y=e^{2x}$, $y=e^{-x}$, $y=2e^{-x}-e^{2x}$, c) $y=e^{-x}-e^{2x}$.
6. $y=C_1+C_2e^{-2x}$, a) $y=1+C_2e^{-2x}$, b) $y=C_1-C_1e^{-2x}$, $C_1<0$.
7. a) $y=C_1+C_2e^x-x^2-2x$, b) $y=C_1e^{2x}+C_2e^{-3x}-5/4e^x$,

c) $y=C_1\cos x+C_2\sin x-2x\cos x+\sin x$, d) $y=(C_1+C_2x)e^x+e^{2x}-1$,
e) $y=C_1e^x+C_2e^{-x}+xe^x-1+(e^{-x}-e^x)\ln(1+e^x)$,
f) $y=(C_1+C_2x-\frac{1}{2}\ln(4+x^2)+\frac{x}{2}\operatorname{arctg}\frac{x}{2})e^x$.

8. a) $y=-xe^{-x}$, b) $y=2x^3+3x^2+9x+3$,
c) $y=(2x-2)e^x$, d) $y=(\frac{1}{2}x^2-\frac{1}{3}x)e^{2x}$,
e) $y=-3x$, f) $y=(x+1)\cos x+x\sin x$,
g) $y=-2xe^{-x}$, h) $y=x+1-e^x$, i) $y=x\sin x$, j) $y=x^2e^{-x}-3\cos x$.

9. a) $y=C_1e^x+C_2e^{-3x}-x$, b) $y=C_1e^{-2x}+C_2xe^{-2x}+\frac{1}{5}(3\cos x+4\sin x)$,
c) $y=(C_1+C_2x+\frac{1}{2}x^3+x^2)e^x$, d) $y=C_1+C_2e^{-2x}-3xe^{-2x}$,
e) $y=(C_1\cos x+C_2\sin x)e^x-x^2+1$, f) $y=C_1+C_2e^x+2x^2+4x+\frac{1}{2}e^{2x}$.

10. a) $y=C_1+C_2e^{-x}-\cos x+\sin x$, b) $y=(C_1\cos 3x+C_2\sin 3x)e^x+\frac{1}{2}x+\frac{1}{2}$,
c) $y=C_1\cos x+C_2\sin x+\frac{\cos^2 x}{\sin x}-\frac{1}{2\sin x}$, d) $y=(C_1+C_2x+\frac{1}{x})e^x$,
e) $y=C_1e^x+C_2e^{-x}-(x^2+x)e^{-x}$, f) $y=(C_1+C_2x+x\ln^2 x-2x\ln x+2x)e^{-x}$,
g) $y=C_1+C_2e^x-xe^x-x$,
h) $y=C_1\cos x+C_2\sin x-\cos x(1+\ln|\sin x|)-x\sin x$,
i) $y=C_1e^{2x}+C_2e^{-2x}-e^x$,
j) $y=C_1\cos x+C_2\sin x-x\cos x+\sin x\ln|\sin x|$,
k) $y=C_1+C_2e^{-2x}+\frac{1}{2}x^3$,
l) $y=C_1e^x+C_2e^{2x}-\frac{1}{5}\cos 2x-\frac{3}{5}\sin 2x+2$.

11. a) $y=(1+2x+x^2)e^{-x}$, b) $y=-3\cos x-3x^2+x+6$,
c) $y=(2+2x\operatorname{arctg}x-\ln(1+x^2))e^{-x}$,
d) $y=\cos x(1+\ln(-\cos x)+x\sin x$, $x\in(\pi/2,3\pi/2)$,
e) $y=\cos x-\sin 2x$, f) $y=2xe^x+(x-1)e^{2x}$.

12. $y=C_1e^x+C_2xe^x+x^2+4x$, $y=e^x+C_2xe^x+x^2+4x$.

14. $y=C_1e^x+C_2e^{-x}+e^{-3x}$, $y=C_1e^x-C_1e^{-x}+e^{-3x}$, $C_1<0$.

15. $y=(C_1+C_2x+x^2)e^x$, a) $y=(x^2+x-5)e^x$,
b) $y=(1+C_2x+x^2)e^x$ dla $C_2\in<-2,2>$.

7. PEWNE WIADOMOŚCI O UKŁADACH RÓWNAŃ RÓŻNICZKOWYCH.

PODSTAWOWE OKREŚLENIA. Niech będzie dany układ równań różniczkowych postaci

$$(7.1)\qquad \begin{cases} \dfrac{dy_1}{dx} = f_1(x,y_1,\ldots,y_n), \\ \ldots\ldots\ldots\ldots\ldots \\ \dfrac{dy_n}{dx} = f_n(x,y_1,\ldots,y_n), \end{cases}$$

gdzie y_k, $k=1,\ldots,n$, oznaczają niewiadome funkcje zmiennej x, a f_k, $k=1,\ldots,n$, są funkcjami ciągłymi na pewnym obszarze $D\subset R^{n+1}$. Jak widać, jest to układ równań różniczkowych rzędu pierwszego, rozwiązanych względem pochodnych niewiadomych funkcji. Układ tej postaci nazywamy **układem normalnym**.

Rozwiązaniem szczególnym układu (7.1) na przedziale I nazywamy każdy układ n funkcji

$$(7.2)\qquad y_1 = y_1(x),\ldots,y_n = y_n(x),\quad x\in I,$$

dla których $\dfrac{dy_k(x)}{dx} = f_k(x,y_1(x),\ldots,y_n(x))$ dla $k=1,\ldots,n$ i każdego $x\in I$, przy czym $(x,y_1(x),\ldots,y_n(x))\in D$ dla dowolnego $x\in I$. Inaczej mówiąc, równania (7.1) dla funkcji (7.2) stają się tożsamościami na przedziale I.

Łatwo na przykład sprawdzić, że para funkcji

$$u = x+e^{2x},\quad v = e^{-2x},\quad x\in R,$$

jest rozwiązaniem szczególnym układu równań

$$\begin{cases} \dfrac{du}{dx} = \dfrac{v+2}{v}, \\ \dfrac{dv}{dx} = \dfrac{2}{x-u}. \end{cases}$$

Układ ten rozważamy przy założeniu, że $u\neq x$ i $v\neq 0$. Podane funkcje u(x) i v(x) warunki te spełniają. Po obliczeniu pochodnych $\dfrac{du}{dx}$, $\dfrac{dv}{dx}$ i wstawieniu do danego układu otrzymujemy równości

$$1+2e^{2x} = \frac{e^{-2x}+2}{e^{-2x}},$$

$$-2e^{-2x} = \frac{2}{-e^{2x}},$$

które są tożsamościami. Zatem para funkcji

$$u = x+e^{2x},\quad v = e^{-2x},\quad x\in R,$$

jest rozwiązaniem danego układu równań.

Rozwiązaniem ogólnym układu (7.1) nazywamy układ n funkcji postaci

$$(7.3)\qquad y_1 = y_1(x,C_1,\ldots,C_n),\ldots,y_n = y_n(x,C_1,\ldots,C_n)$$

($C_1,\ldots,C_n$ oznaczają tu dowolne stałe) mający tę własność, że dla dowolnego dopuszczalnego układu stałych $C_1,\ldots,C_n$ funkcje (7.3) stanowią pewne rozwiązanie układu (7.1).

Zagadnienie Cauchy'ego dla układu równań różniczkowych (7.1) jest to zadanie polegające na znalezieniu takiego rozwiązania

$$y_1 = y_1(x),\ldots,y_n = y_n(x)$$

tego układu, które spełnia warunki

$$y_1(x_0) = y_1^0,\ldots,y_n(x_0) = y_n^0$$

dla danych liczb $x_0,y_1^0,\ldots,y_n^0$.

W dalszych rozważaniach ograniczymy się do układów w postaci normalnej. Warto jednak zauważyć, że ograniczenie to nie jest bardzo istotne. Wiele układów równań zapisanych w innej postaci można sprowadzić do układu normalnego. Na przykład łatwo wykazać, że każdy układ równań rozwiązanych ze względu na najwyższe pochodne niewiadomych funkcji (tzw. układ standardowy) można sprowadzić do układu normalnego poprzez wprowadzenie pomocniczych niewiadomych. Ilustruje to następujący przykład.

PRZYKŁAD 7.1 Niech będzie dany układ równań

$$\begin{cases} \dfrac{d^2y}{dx^2} = y + z - \dfrac{dy}{dx}, \\ \dfrac{dz}{dx} = 3x + z^2 + 2\dfrac{dy}{dx}. \end{cases}$$

Jest to układ dwóch równań z niewiadomymi funkcjami $y(x)$ i $z(x)$, przy czym równania te są rozwiązane ze względu na najwyższe pochodne niewiadomych funkcji występujące w tym układzie, tzn. y'' i z'. Wprowadźmy nowe niewiadome funkcje u, v, w zmiennej x, za pomocą równości: $u = y$, $v = \dfrac{dy}{dx}$, $w = z$. Wtedy $\dfrac{d^2y}{dx^2} = \dfrac{dv}{dx}$. Po uwzględnieniu wprowadzonych oznaczeń, otrzymujemy układ w postaci normalnej z funkcjami niewiadomymi $u(x)$, $v(x)$ i $w(x)$:

$$\begin{cases} \dfrac{du}{dx} = v, \\ \dfrac{dv}{dx} = u + w - v, \\ \dfrac{dw}{dx} = 3x + w^2 + 2v. \end{cases}$$

■

Zauważmy jeszcze, że każde równanie różniczkowe rzędu n postaci

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)}), \tag{7.4}$$

gdzie $y = y(x)$ jest niewiadomą funkcją, można sprowadzić do układu normalnego n równań wprowadzając pomocnicze funkcje niewiadome $u_k = u_k(x)$, $k = 1, \ldots, n$:

$$u_1 = y, \quad u_2 = y', \ldots, u_n = y^{(n-1)}.$$

W konsekwencji otrzymujemy układ równań

$$\begin{cases} \dfrac{du_1}{dx} = u_2, \\ \dfrac{du_2}{dx} = u_3, \\ \ldots\ldots\ldots\ldots \\ \dfrac{du_n}{dx} = f(x, u_1, \ldots, u_n), \end{cases} \tag{7.5}$$

który jest układem normalnym z niewiadomymi funkcjami $u_1(x), \ldots, u_n(x)$.

Stąd wynika, że badanie rozwiązalności równania różniczkowego rzędu n można zastąpić badaniem rozwiązalności normalnego układu n równań różniczkowych. I rzeczywiście, dowód istnienia rozwiązań równania postaci (7.4) przeprowadza się w ten sposób, że wykazuje się istnienie rozwiązań odpowiadającego mu układu równań (7.5).

Niżej podajemy **twierdzenie o istnieniu i jednoznaczności rozwiązań układu równań różniczkowych** w postaci normalnej (7.1).

TWIERDZENIE 7.1. Jeżeli funkcje f_k, $k = 1, \ldots, n$, są ciągłe i mają ciągłe pochodne cząstkowe względem $y_1, \ldots, y_n$ na pewnym obszarze $D \subset R^{n+1}$ oraz $(x_0, y_1^0, \ldots, y_n^0) \in D$, to istnieje dokładnie jedno rozwiązanie układu (7.1) określone na pewnym otoczeniu punktu x_0 i spełniające warunki początkowe

$$y_1(x_0) = y_1^0, \ldots, y_n(x_0) = y_n^0.$$

Łatwo sprawdzić na przykład, że zgodnie z tym twierdzeniem dla dowolnych liczb x_0, y_0, z_0 istnieje dokładnie jedno rozwiązanie układu

$$\begin{cases} \dfrac{dy}{dx} = y^2 - z + \cos x, \\ \dfrac{dz}{dx} = 2y^3 + z^2 - x^3 \end{cases}$$

określone na pewnym otoczeniu punktu x_0 i spełniające warunki $y(x_0) = y_0$ i $z(x_0) = z_0$. Nie oznacza to jednak, że rozwiązanie takie potrafimy łatwo wyznaczyć. Podobnie jak dla równań różniczkowych, tylko dla pewnych układów równań (np. dla układów równań liniowych) potrafimy wskazać efektywne metody całkowania, czyli znajdowania rozwiązań. Na ogół możemy jedynie zastosować pewne metody przybliżonego rozwiązywania.

Pokażemy teraz jedną z metod rozwiązywania układów równań różniczkowych polegającą na sprowadzeniu układu do jednego równania różniczkowego.

METODA ELIMINACJI rozwiązywania układu. Na początku tego paragrafu pokazaliśmy m.in. sposób sprowadzenia równania rzędu n do układu normalnego n równań. Przy rozwiązywaniu układów

czasem okazuje się celowe postępowanie odwrotne - sprowadzenie układu równań, poprzez rugowanie kolejnych niewiadomych funkcji, do jednego równania odpowiednio wysokiego rzędu z jedną niewiadomą funkcją. Rozwiązanie tego równania pozwala nam wyznaczyć pozostałe niewiadome funkcje już bez całkowania. Metodę eliminacji zilustrujemy przykładami.

PRZYKŁAD 7.2. Rozwiążemy układ równań

$$(1)\qquad \begin{cases} y' = u, \\ z' = u + 2e^x, \\ u' = z - y - e^x, \end{cases}$$

w którym niewiadome funkcje y, z i u są funkcjami zmiennej x.

Jest to układ normalny trzech równań. Eliminując dwie niewiadome funkcje sprowadzimy ten układ do jednego równania różniczkowego rzędu trzeciego z jedną funkcją niewiadomą. Z pierwszego równania układu (1) mamy $u = y'$, a zatem $u' = y''$. Uwzględniając to w pozostałych równaniach mamy

$$(2)\qquad \begin{cases} z' = y' + 2e^x, \\ y'' = z - y - e^x \end{cases}$$

(w ten sposób wyeliminowana została niewiadoma u). Następnie rugujemy z otrzymanego układu (2) funkcję z. W tym celu z drugiego równania układu (2) wyznaczymy z:

$$z = y'' + y + e^x,$$

obliczamy pochodną z':

$$z' = y''' + y' + e^x$$

i wstawiamy do pierwszego z równań układu (2). Otrzymujemy

$$y''' = e^x.$$

Jest to równanie różniczkowe rzędu trzeciego z niewiadomą funkcją y. Przez trzykrotne całkowanie łatwo otrzymujemy rozwiązanie ogólne równania:

$$y = C_1x^2 + C_2x + C_3 + e^x,$$

gdzie C_1, C_2, C_3 oznaczają dowolne stałe. Ponieważ $u = y'$ oraz $z = y'' + y + e^x$, więc

$$u = 2C_1x + C_2 + e^x$$

oraz

$$z = C_1x^2 + C_2x + +2C_1 + C_3 + 3e^x.$$

Rozwiązaniem ogólnym układu (1) jest układ funkcji

$$(3)\qquad \begin{aligned} y &= C_1x^2 + C_2x + C_3 + e^x, \\ z &= C_1x^2 + C_2x + 2C_1 + C_3 + 3e^x, \\ u &= 2C_1x + C_2 + e^x, \end{aligned}$$

gdzie $C_1, C_2, C_3 \in R$ i $x \in R$.

Zgodnie z podanym wcześniej twierdzeniem 7.1 o istnieniu i jednoznaczności rozwiązań wiadomo, że dla dowolnej czwórki liczb (x_0, y_0, z_0, u_0) istnieje dokładnie jedno rozwiązanie układu równań różniczkowych (1) spełniające warunki początkowe:

$$y(x_0) = y_0, \quad z(x_0) = z_0, \quad u(x_0) = u_0.$$

Wyznaczmy więc takie rozwiązanie szczególne układu (1), które spełnia warunki $y(0) = 1$, $z(0) = 3$, $u(0) = -1$. Uwzględniając podane warunki początkowe w (3) otrzymujemy

$$\begin{aligned} y(0) &= C_3 + 1 = 1, \\ z(0) &= 2C_1 + C_3 + 3 = 3, \\ u(0) &= C_2 + 1 = -1, \end{aligned}$$

skąd

$$C_1 = 0, \qquad C_2 = -2, \qquad C_3 = 0.$$

Rozwiązaniem układu (1) spełniającym postawione warunki jest więc następujący układ funkcji:

$$y = e^x - 2x, \quad z = 3e^x - 2x, \quad u = e^x - 2, \quad x \in R. \qquad \blacksquare$$

PRZYKŁAD 7.3. Rozwiążemy układ równań

$$(1)\qquad \begin{cases} \dfrac{dx}{dt} = x + 4y + 12e^t, \\ \dfrac{dy}{dt} = x + y - 3e^t. \end{cases}$$

Jest to układ dwóch równań z niewiadomymi funkcjami x(t) i y(t). Z drugiego równania wyznaczymy x:

$$x = y' - y + 3e^t.$$

Po zróżniczkowaniu i wstawieniu do pierwszego równania mamy

$$y'' - y' + 3e^t = y' - y + 3e^t + 4y + 12e^t,$$

czyli

$$y'' - 2y' - 3y = 12e^t. \tag{2}$$

Jest to równanie liniowe drugiego rzędu o stałych współczynnikach z niewiadomą funkcją $y = y(t)$. Rozwiązaniem ogólnym równania jednorodnego $y'' - 2y' - 3y = 0$ jest $y_0 = C_1e^{-t} + C_2e^{3t}$. Metodą przewidywania łatwo znajdujemy rozwiązanie szczególne równania (2): $y_s = -3e^t$. Zatem rozwiązanie ogólne równania (2) ma postać

$$y = C_1e^{-t} + C_2e^{3t} - 3e^t.$$

Ponieważ $x = y' - y + 3e^t$, więc

$$x = -2C_1e^{-t} + 2C_2e^{3t} + 3e^t.$$

Rozwiązanie ogólne układu (1) ma postać

$$x = -2C_1e^{-t} + 2C_2e^{3t} + 3e^t, \quad y = C_1e^{-t} + C_2e^{3t} - 3e^t, \quad t \in R,$$

gdzie C_1, C_2 oznaczają dowolne stałe. ■

PRZYKŁAD 7.4. Rozwiążemy układ równań

$$\begin{cases} \dfrac{dy}{dx} = \dfrac{y(y-1)}{z}, \\ \dfrac{dz}{dx} = y - 1 \end{cases} \tag{1}$$

przy warunkach początkowych $y(0) = 2$, $z(0) = 1$.

Z drugiego równania wynika, że $y = z' + 1$, a stąd $y' = z''$. Po podstawieniu do pierwszego równania otrzymujemy

$$z'' = \frac{(z'+1)z'}{z}. \tag{2}$$

Jest to równanie rzędu drugiego, w którym nie występuje w sposób wyraźny zmienna x. Równanie takie rozwiązujemy przez wprowadzenie pomocniczej niewiadomej. Niech

$$z' = u(z).$$

Wówczas

$$z'' \equiv \frac{d^2z}{dx^2} = \frac{du}{dz}\frac{dz}{dx} = \frac{du}{dz}u.$$

Po wykonaniu tego podstawienia w równaniu (2) otrzymujemy równanie I rzędu o zmiennych rozdzielonych:

$$\frac{du}{dz}u = \frac{(u+1)u}{z}. \tag{3}$$

Rozwiązaniami równania (3) są funkcje $u = u(z)$ postaci

$$u = C_1z - 1, \ C_1 \in R, \ z \neq 0; \qquad u = 0, \ z \neq 0.$$

Następnie uwzględniając, że $z'(x) = u(z)$, otrzymujemy rozwiązania $z = z(x)$ równania (2):

$$z = \frac{1}{C_1}(1 + C_2e^{C_1x}), \ C_1 \neq 0, \ C_2 \in R, \ z \neq 0; \quad z = -x + C, \ C \in R, \ z \neq 0.$$

Ponieważ $y = z' + 1$, więc rozwiązaniami układu (1) są pary funkcji

$$y = 1 + C_2e^{C_1x}, \quad z = \frac{1}{C_1}(1 + C_2e^{C_1x}), \ C_1 \neq 0, \ C_2 \in R, \tag{4}$$

oraz

$$y = 0, \quad z = -x + C, C \in R, \tag{5}$$

przy czym x jest takie, że $z \neq 0$.

Pozostaje jeszcze spośród wszystkich tych rozwiązań wybrać to, które czyni zadość postawionym warunkom. Łatwo sprawdzić, że żadna para funkcji określona wzorami (5) warunków tych nie spełnia. Natomiast uwzględniając warunki początkowe w (4) otrzymujemy

$$y(0) = 1 + C_2 = 2,$$

$$z(0) = \frac{1}{C_1}(1 + C_2) = 1,$$

a stąd $C_1 = 2$, $C_2 = 1$. Szukanym rozwiązaniem szczególnym układu równań (1) jest para funkcji:

$$y = 1 + e^{2x}, \quad z = \frac{1}{2}(1 + e^{2x}), \quad x \in R. \qquad ■$$

Przedstawiona tu metoda eliminacji rozwiązywania układów równań polega na zastąpieniu układu równań jednym równaniem odpowiednio wyższego rzędu. Należy jednak zauważyć, że nie każdy układ normalny można sprowadzić do jednego równania, o czym przekona nas następujący prosty przykład.

PRZYKŁAD 7.5. Rozwiążemy układ równań

$$(1)\qquad \begin{cases} \dfrac{dx}{dt} = \dfrac{2(x-1)}{t}, \\ \dfrac{dy}{dt} = y - 4e^{-t}. \end{cases}$$

Jest to układ dwóch równań niezależnych - każde z tych równań jest równaniem różniczkowym rzędu pierwszego z jedną niewiadomą funkcją zmiennej t. W pierwszym równaniu nie występuje funkcja y(t), a w drugim nie występuje funkcja x(t). Takiego układu nie można zastąpić jednym równaniem drugiego rzędu z jedną niewiadomą funkcją. Oczywiście nie jest to konieczne dla rozwiązania tego układu. Rozwiązaniem układu (1) jest każda para funkcji x i y, z których pierwsza jest rozwiązaniem pierwszego równania , zaś druga - rozwiązaniem drugiego równania. Z pierwszego równania układu (1) otrzymujemy $x = 1 + C_1 t^2$ dla $t \neq 0$, a z drugiego $y = C_2 e^t + 2e^{-t}$, gdzie C_1 i C_2 są dowolnymi stałymi. Zatem układ funkcji

$$x = 1 + C_1 t^2, \qquad y = C_2 e^t + 2e^{-t}, \quad t \neq 0,$$

gdzie C_1 i C_2 oznaczają dowolne stałe, jest rozwiązaniem ogólnym rozważanego układu równań. ■

ZADANIA DO ROZWIĄZANIA.

1. Rozwiązać układ równań:

a) $\begin{cases} \dfrac{dy}{dx} = 2y + z, \\ \dfrac{dz}{dx} = y + 2z, \end{cases}$ b) $\begin{cases} \dfrac{dy}{dx} = y - 5z, \\ \dfrac{dz}{dx} = y - z, \end{cases}$

c) $\begin{cases} \dfrac{dy}{dx} = y + 3z - 8x^2, \\ \dfrac{dz}{dx} = y - z, \end{cases}$ d) $\begin{cases} \dfrac{dx}{dt} = y + x, \\ \dfrac{dy}{dt} = 2y + 6e^{-2t}, \end{cases}$

e) $\begin{cases} \dfrac{dy}{dt} + y + 2z = 3t^2, \\ \dfrac{dz}{dt} - y - 2z = -6, \end{cases}$ f) $\begin{cases} \dfrac{dx}{dt} - 2x - y = e^t, \\ \dfrac{dy}{dt} + 2y + 5x = -e^t, \end{cases}$

g) $\begin{cases} \dfrac{dx}{dt} = \dfrac{x^2}{y}, \\ \dfrac{dy}{dt} = x, \end{cases}$ h) $\begin{cases} \dfrac{dy}{dx} = z + 1, \\ \dfrac{dz}{dx} = \dfrac{z}{x} - \dfrac{2}{x^2}. \end{cases}$

2. Rozwiązać układ równań przy podanych warunkach początkowych:

a) $\begin{cases} \dfrac{dy}{dx} = 2y + z, \\ \dfrac{dz}{dx} = 4y - z, \end{cases}$ $y(0) = -1,\ z(0) = 4;$

b) $\begin{cases} \dfrac{dy}{dx} - 3y + 4z = 0, \\ \dfrac{dz}{dx} - y + z = 0, \end{cases}$ $y(1) = e,\ z(1) = 0;$

c) $\begin{cases} \dfrac{dx}{dt} = -1 - x - 2y, \\ \dfrac{dy}{dt} = t + x + y, \end{cases}$ $x(0) = 2,\ y(0) = -1.$

Odpowiedzi. Litery C_1 i C_2 oznaczają dowolne stałe

1. a) $y = C_1 e^x + C_2 e^{3x}$, $z = -C_1 e^x + C_2 e^{3x}$,
   b) $y = (C_2 - 2C_1)\sin 2x + (2C_2 + C_1)\cos 2x$, $z = C_1 \cos 2x + C_2 \sin 2x$
   c) $y = 3C_1 e^{2x} - C_2 e^{-2x} + 2x^2 + 4x + 1$, $z = C_1 e^{2x} + C_2 e^{-2x} + 2x^2 + 1$,
   d) $x = C_1 e^t + C_2 e^{2t} + \frac{1}{2} e^{-2t}$, $y = C_2 e^{2t} - \frac{3}{2} e^{-2t}$,
   e) $y = -2C_1 - C_2 e^t + 2t^3 + 3t^2 - 6t + 6$, $z = C_1 + C_2 e^t - t^3 - 3t^2$,
   f) $x = C_1 \cos t + C_2 \sin t + e^t$, $y = (C_2 - 2C_1)\cos t - (C_1 + 2C_2)\sin t - 2e^t$,
   g) $x = C_1 C_2 e^{C_1 t}$, $y = C_2 e^{C_1 t}$, $C_2 \neq 0$, h) $y = x + \ln|x| + \frac{C_1}{2} x^2 + C_2$, $z = \frac{1}{x} + C_1 x$.
2. a) $y = -e^{-2x}$, $z = 4e^{-2x}$, b) $y = (2x-1)e^x$, $z = (x-1)e^x$,
   c) $x = \sin t + \cos t - 2t + 1$, $y = t - \cos t$.

V. ELEMENTY RACHUNKU PRAWDOPODOBIEŃSTWA

W tym i następnym rozdziale będziemy stosować następujące skróty:
PZE - przestrzeń zdarzeń elementarnych,
pr-stwo - prawdopodobieństwo,
ZL - zmienna losowa,
ZLS - zmienna losowa skokowa,
ZLC - zmienna losowa ciągła,
GP - gęstość pr-stwa,
WL - wektor losowy,
WLS - wektor losowy skokowy,
WLC - wektor losowy ciągły,
PLP - próba losowa prosta.

1. ZDARZENIA LOSOWE.

DOŚWIADCZENIE LOSOWE. Doświadczenie losowe (krócej: **doświadczenie**) - to realizacja (rzeczywista lub tylko myślowa) zespołu warunków wraz ze zbiorem obserwowanych wyników. Tak więc termin ten będziemy rozumieć w bardzo szerokim sensie: zarówno z góry zaplanowany eksperyment jak i niezaplanowana obserwacja są doświadczeniami losowymi.

PRZYKŁAD 1.1. Podajemy kilka przykładów doświadczeń:
D_1 - rzut monetą i obserwacja górnej strony monety,
D_2 - rzut monetą i obserwacja czy będzie ona spadać,

D_3 - zdawanie egzaminu i obserwacja wyniku, jeśli egzamin kończy się wpisem do indeksu: "zaliczono", "nie zaliczono",
D_4 - zdawanie egzaminu i obserwacja uzyskanego wyniku przy sześciostopniowej skali ocen: 1, 2, 3, 4, 5, 6,
D_5 - rzut monetą aż do zaobserwowania orła po raz pierwszy, jednak nie więcej niż 3 razy,
D_6 - rzut monetą aż do zaobserwowania orła po raz pierwszy,
D_7 - obserwacja czasu bezawaryjnej pracy pewnego elementu danego urządzenia,
D_8 - badanie wytrzymałości na zrywanie odcinka liny stalowej i notowanie zaobserwowanej wartości,
D_9 - badanie wytrzymałości na zrywanie odcinka liny stalowej i obserwacja czy spełnia on wymagania normy. ■

Podkreślamy, że dwa doświadczenia są różne również wtedy, gdy stanowią realizację tych samych warunków, ale różnią się zbiorem obserwowanych wyników. Na przykład różne są pary doświadczeń: D_1 i D_2, D_3 i D_4, D_5 i D_6, D_8 i D_9.

Wśród warunków, których realizacja stanowi doświadczenie można wyróżnić: **warunki systematyczne (zdeterminowane, główne)** - dające się zadawać, regulować, kontrolować oraz **warunki losowe** - wymykające się naszej kontroli, nie dające się regulować. Udział poszczególnych rodzajów warunków bywa różny; w szczególności jedne z nich mogą w ogóle nie wystąpić. Chętnie podaje się jako przykłady doświadczeń losowych takie doświadczenia, w których nie występują warunki systematyczne; takimi są na przykład doświadczenia D_1, D_5, D_6. Z kolei w doświadczeniu D_2 nie występują warunki losowe.

PRZESTRZEŃ ZDARZEŃ ELEMENTARNYCH. Częścią składową doświadczenia jest zbiór jego wyników. Mamy tu na myśli wyniki najprostsze, z których dokładnie jeden realizuje się w rezultacie przeprowadzenia doświadczenia. Te najprostsze wyniki doświadczenia nazywa się wynikami elementarnymi albo **zdarzeniami elementarnymi** doświadczenia, a ich zbiór - **przestrzenią zdarzeń elementarnych (PZE)** i oznacza literą Ω. Elementy PZE Ω oznaczamy małą literą ω, $\omega \in \Omega$.

PRZYKŁAD 1.2. Określimy teraz PZE Ω_i dla doświadczeń D_i z przykładu 1.1:

Ω_1 = {orzeł, reszka} ≡ {O,R}, Ω_2 = {spadanie monety},
Ω_3 = {zaliczono, nie zaliczono}, Ω_4 = {1, 2, 3, 4, 5, 6},
Ω_5 = {O, RO, RRO, RRR},
Ω_6 = {O, RO, RRO, RRRO, RRRRO, RRRRRO, ...},
Ω_7 = $<0,t>$, $t \in R$, $\Omega_8 = (0,\infty)$,
Ω_9 = {spełnia wymagania, nie spełnia wymagań}. ■

Analiza tego przykładu prowadzi do następujących spostrzeżeń.

1) Zdarzenia elementarne ω mogą mieć bardzo różnorodną naturę,

2) PZE Ω może być zbiorem skończonym (tak jak przestrzenie $\Omega_1,\Omega_2,\Omega_3,\Omega_4,\Omega_5,\Omega_9$), nieskończonym, ale przeliczalnym (jak Ω_6), nieskończonym i nieprzeliczalnym (jak Ω_7, Ω_8). PZE skończoną lub nieskończoną, ale przeliczalną (zbiór nazywa się **zbiorem przeliczalnym**, gdy jego elementy dają się ustawić w ciąg nieskończony) nazywamy **skokową** (ziarnistą). Zbiory takie jak Ω_7, Ω_8 nazywa się **ciągłymi** PZE. Skokowe i ciągłe PZE nie wyczerpują wszystkich możliwych rodzajów PZE. Przypomnijmy, że w szkolnym kursie rachunku pr-stwa rozważane są tylko skończone PZE.

W jedynie poprawnym aksjomatycznym ujęciu rachunku pr-stwa zdarzenia elementarne ω i PZE Ω są **pojęciami pierwotnymi** (tj. takimi, których się nie definiuje). W każdym doświadczeniu, jeśli w jego opracowaniu stosujemy rachunek pr-stwa, ustala się co jest PZE; zwykle można to uczynić na wiele sposobów.

ZDARZENIA LOSOWE. Często ważniejsze są nie pojedyncze zdarzenia elementarne ω, lecz ich zbiory. Na przykład nie jest najważniejsze, aby wytrzymałość X liny stalowej była równa dokładnie x_1, $X=x_1$, lecz aby $X \geq x_1$. P o d z b i o r y PZE Ω nazywa się **zdarzeniami losowymi** (krócej: **zdarzeniami**). Będziemy je oznaczać (tak jak zbiory) dużymi literami A, B, C ... . Zdarzenia elementarne ω, z których składa się zdarzenie A nazywa się zdarzeniami elementarnymi **sprzyjającymi** zajściu zdarzenia A.

Dodajmy, że w pełnym wykładzie rachunku pr-stwa nie zawsze wszystkie podzbiory przestrzeni Ω przyjmuje się jako zdarzenie losowe.

Zbiór wszystkich zdarzeń oznaczymy stylizowaną literą $\mathcal{A}$; zatem elementami zbioru $\mathcal{A}$ są zdarzenia A, B, C, ...

P R Z Y K Ł A D 1.3. Podajemy przykłady zdarzeń związanych odpowiednio z doświadczeniami D_4, D_5, D_6, D_8 z przykładów 1.1 i 1.2:

A = zdanie egzaminu = {2,3,4,5,6}, $A \subset \Omega_4$,

A = orzeł pojawi się w rzucie o numerze nieparzystym =
= {O, RRO}, $A \subset \Omega_5$,

A = orzeł pojawi się w rzucie o numerze nieparzystym =
= {O, RRO, RRRRO, ... }, $A \subset \Omega_6$,

A = wytrzymałość przyjmuje wartość równą co najmniej x_1 i jednocześnie nie większą od x_2 = przedział domknięty $<x_1,x_2>$, $A \subset \Omega_8$. ■

Zdarzenia są zbiorami. Zatem między zdarzeniami zachodzą takie same relacje jak między zbiorami; w szczególności na zdarzeniach wykonuje się takie same działania jak na zbiorach; jednak niektóre nazwy są inne. Krótko przypomnimy je Czytelnikowi.

Zdarzenie pewne jest to cała PZE Ω.

Zdarzenie niemożliwe jest to podzbiór pusty $\emptyset$ zbioru Ω.

Koniunkcja (iloczyn) $A_1 \cap A_2 \cap \cdots$ **zdarzeń** $A_1, A_2, \ldots$ jest to zdarzenie składające się z tych zdarzeń elementarnych ω, które sprzyjają w s z y s t k i m zdarzeniom A_i. Jest to więc zdarzenie polegające na zajściu w s z y s t k i c h zdarzeń A_i

Alternatywa (suma) $A_1 \cup A_2 \cup \cdots$ **zdarzeń** $A_1, A_2, \ldots$ jest to zdarzenie składające się z tych zdarzeń elementarnych ω, które sprzyjają c o n a j m n i e j jednemu ze zdarzeń A_i. Jest to więc zdarzenie polegające na zajściu c o n a j m n i e j jednego ze zdarzeń A_i.

Różnica $B \setminus A$ **zdarzeń** B, A jest to zdarzenie składające się z tych zdarzeń elementarnych ω, które sprzyjają zdarzeniu B i nie sprzyjają zdarzeniu A. Jest to więc zdarzenie polegające na zajściu zdarzenia B i niezajściu zdarzenia A.

Mówimy, że zdarzenia A, B są **rozłączne (wykluczają się)**, jeśli nie mają wspólnych zdarzeń elementarnych ω, tj. gdy $A \cap B = \emptyset$.

Zdarzenia A' i A nazywa się **przeciwnymi**, jeśli $A' = \Omega \setminus A$. Definicja ta jest równoważna następującej: mówimy, że zdarzenia A i A' są przeciwne, jeśli 1) $A \cap A' = \emptyset$ (są rozłączne) oraz 2) $A \cup A' = \Omega$ (ich alternatywa jest zdarzeniem pewnym).

Mówimy, że zdarzenia $A_1, A_2, \ldots$ są **parami rozłączne**, jeśli każde dwa z nich są rozłączne.

Mówimy, że zdarzenia $A_1, A_2, \ldots$ tworzą **układ zupełny zdarzeń**, jeśli: 1) są parami rozłączne, 2) ich alternatywa jest zdarzeniem pewnym. Na przykład zdarzenia A, A' tworzą układ zupełny zdarzeń.

Mówimy, że zdarzenie A **pociąga** zdarzenie B, co zapisujemy $A \subset B$, jeśli każde zdarzenie elementarne ω sprzyjające zdarzeniu A sprzyja również zdarzeniu B.

Zdarzenia A i B są **równe**, $A = B$, jeśli $A \subset B$ oraz $B \subset A$.

Przypominamy jeszcze niektóre własności działań na zdarzeniach (są to własności działań na zbiorach):

$$\left.\begin{array}{l} A \cap B = B \cap A \\ A \cup B = B \cup A \end{array}\right\}$$ przemienność koniunkcji i alternatywy zdarzeń,

$$\left.\begin{array}{l} A \cap (B \cap C) = (A \cap B) \cap C \\ A \cup (B \cup C) = (A \cup B) \cup C \end{array}\right\}$$ łączność koniunkcji i alternatywy zdarzeń,

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ - rozdzielność koniunkcji wzgl. alternatywy,

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ - rozdzielność alternatywy wzgl. koniunkcji,

$$\left.\begin{array}{l} (A \cap B)' = A' \cup B' \\ (A \cup B)' = A' \cap B' \end{array}\right\}$$ prawa de Morgana.

ZADANIA DO ROZWIĄZANIA.

1. Zaproponować PZE Ω dla następujących doświadczeń D_1, D_2, D_3, D_4, D_5
   - D_1 - lokalizowanie awarii zespołu trzech maszyn, jeśli wiadomo, że może ona być rezultatem awarii jednej, dwu lub wszystkich trzech maszyn,
   - D_2 - obserwacja liczby zgłoszeń do centrali telefonicznej w ciągu 8 godzin,
   - D_3 - powtarzanie doświadczenia (np. losowania ze zwrotem) aż do otrzymania interesującego nas wyniku,
   - D_4 - wbijanie pali o losowej długości L nie większej od l_{max} na losową głębokość H warstwy skalnej, położonej nie niżej niż h_{max},
   - D_5 - obserwacja krzywych elektrokardiogramu u pacjentów.

2. Niech A, B, C będą dowolnymi zdarzeniami. Za pomocą działań na zdarzeniach opisać zdarzenia polegające na zajściu spośród tych zdarzeń: a) wszystkich trzech zdarzeń, b) co najmniej jednego z tych zdarzeń, c) dokładnie jednego z tych zdarzeń.
3. Niech D_1 będzie doświadczeniem D_1 z zadania 1. Niech A, B, C oznaczają odpowiednio zdarzenia: pierwsza, druga, trzecia sprawdzana maszyna jest uszkodzona. Za pomocą zdarzeń elementarnych tego doświadczenia opisać zdarzenia określone w punktach a, b, c zadania 2.
4. Uzupełnić równości: a) $A \cap A =$, b) $A \cup A =$, c) $A \cap \Omega =$, d) $A \cup \Omega =$.
5. $A \subset B$, uzupełnić równości: a) $A \cap B =$, b) $A \cup B =$, c) $A \cap B \cap C =$
6. Uprościć opis zdarzenia A, jeśli: a) $A = (B \cup C) \cap (B' \cup C)$,
   b) $A = (B \cup C) \cap (B' \cup C) \cap (B \cup C')$,
   c) $A = (B \cup C) \cap (C \cup E)$, gdy $B \subset E$.

Odpowiedzi.

1. $\Omega_1 = \{111, 110, 101, 011, 100, 010, 001, 000\}$, gdzie 1 na i-tym miejscu w trzywyrazowym ciągu oznacza, że i-ta maszyna pracowała niezawodnie oraz 0 na i-tym miejscu informuje, że i-ta maszyna uległa awarii; $\Omega_2 = \{0,1,2,\ldots,n\}$; $\Omega_3 = \{1,2,\ldots,n,\ldots\}$; $\Omega_4 = \{(L,H): 0 < L \le l_{max}, 0 \le H \le h_{max}\}$ = zbiór punktów prostokąta bez jednego boku; Ω_5 = zbiór wykresów funkcji jednej zmiennej.
2. a) $A \cap B \cap C$; b) $A \cup B \cup C$; c) $(A \cap B' \cap C') \cup (A' \cap B \cap C') \cup (A' \cap B' \cap C)$.
3. a) $\{000\}$; b) $\{110, 101, 011, 100, 010, 001, 000\}$; c) $\{110, 101, 011\}$.
4. a) A; b) A; c) A; d) Ω. 5. a) A; b) B; c) $A \cap C$. 6. a) C; b) $B \cap C$; c) $B \cup C$.

2. PRAWDOPODOBIEŃSTWO.

AKSJOMATYCZNA DEFINICJA PRAWDOPODOBIEŃSTWA.

Przed przeprowadzeniem doświadczenia nie można rozstrzygnąć, czy zajdzie związane z nim i interesujące nas zdarzenie losowe A (właśnie dlatego, że jest losowe). Można natomiast mówić o ocenie szansy zajścia tego zdarzenia; miarą tej szansy jest prawdopodobieństwo (pr-stwo) zajścia zdarzenia (krócej: pr-stwo zdarzenia). Pojęcie to, znane Czytelnikowi w przypadku skończonej PZE, rozszerzymy na przypadek dowolnej skokowej lub ciągłej PZE.

W zrozumieniu definicji pr-stwa pomaga uświadomienie sobie, że w definicji funkcji nie zakłada się nic o naturze elementów zbioru stanowiącego jej dziedzinę. W szczególności elementy te mogą być zbiorami (np. dziedziną może być zbiór prostokątów, wartościami funkcji - pola tych prostokątów). Funkcję, która zbiorom przyporządkowuje liczby nazywa się **funkcją zbioru**. Pr-stwo jest taką właśnie funkcją zbioru, która zdarzeniom (a więc zbiorom) przyporządkowuje pr-stwa tych zdarzeń.

Niech zatem Ω oznacza PZE oraz $\mathcal{A}$ będzie zbiorem zdarzeń, czyli zbiorem (rodziną) podzbiorów przestrzeni Ω. **Prawdopodobieństwem** (lub **rozkładem prawdopodobieństwa**) nazywamy funkcję P, która każdemu zdarzeniu $A \in \mathcal{A}$ przyporządkowuje liczbę P(A) tak, aby spełnione zostały następujące warunki:

P1° $P(A) \geq 0$ dla każdego zdarzenia A,

P2° $P(\Omega) = 1$,

P3° dla dowolnego ciągu parami rozłącznych zdarzeń $A_1, A_2, \ldots$

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

Wartość pr-stwa P dla danego zdarzenia A, czyli liczbę P(A) nazywa się **pr-stwem zdarzenia** A.

Warunki P1°, P2°, P3° nazywa się **aksjomatami pr-stwa**, a samą definicję - **aksjomatyczną definicją pr-stwa**. Jako taka, podaje warunki, jakie winna spełniać funkcja zbioru, aby można było nazwać ją pr-stwem. Nie precyzuje ona konkretnych wartości pr-stwa dla poszczególnych zdarzeń (z wyjątkiem zdarzenia pewnego). Konkretyzacja taka jest możliwa tylko dla poszczególnych doświadczeń przez uzupełnienie aksjomatów dodatkowymi warunkami charakterystycznymi dla rozważanego doświadczenia. Przykładem takiego dodatkowego warunku może być na przykład założenie, że zdarzenia jednoelementowe są jednakowo prawdopodobne.

Własności prawdopodobieństwa. Ze związków między zdarzeniami i definicji pr-stwa wynikają dalsze własności tego pojęcia sformułowane w następującym twierdzeniu; ich treścią nie są (z jednym wyjątkiem) konkretne wartości pr-stwa poszczególnych zdarzeń, ale wzajemne zależności między nimi.

TWIERDZENIE 2.1. Pr-stwo P ma następujące własności:

P1. Pr-stwo zdarzenia niemożliwego jest równe zero: $P(\emptyset) = 0$.

P2. Dla dowolnych zdarzeń A, B pr-stwo różnicy $B \setminus A$ wyraża się wzorem:

$$P(B \setminus A) = P(B) - P(A \cap B);$$

w szczególności, gdy $A \subset B$, to

$$P(B \setminus A) = P(B) - P(A).$$

P3. Pr-stwo P jest funkcją niemalejącą w następującym sensie:

$$A \subset B \Rightarrow P(A) \leq P(B).$$

P4. Pr-stwo P jest funkcją o wartościach z przedziału domkniętego $\langle 0;1 \rangle$, czyli dla każdego zdarzenia A

$$0 \leq P(A) \leq 1.$$

P5. Jeżeli zdarzenia $A_1, A_2, \ldots, A_n$ są parami rozłączne, to

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$

P6. Jeżeli zdarzenia $A_1, A_2, \ldots, A_n$ tworzą układ zupełny zdarzeń, to

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1.$$

P6'. Suma pr-stw zdarzeń przeciwnych jest równa jedności:

$$P(A) + P(A') = 1.$$

P7. Pr-stwo alternatywy dowolnych zdarzeń A, B wyraża się równością:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

P7'. Dla dowolnych zdarzeń A, B, C

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

P8. Dla dowolnych zdarzeń $A_1, A_2, \ldots, A_n$

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) \leq P(A_1) + P(A_2) + \cdots + P(A_n).$$

P9. Jeżeli: 1) PZE Ω jest skończona $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ albo przeliczalna $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$ oraz 2) określone są pr-stwa α_i zdarzeń jednoelementowych $P(\{\omega_i\}) = \alpha_i$, $i = 1,2,\ldots,n$ lub $i = 1,2,\ldots,n,\ldots$, gdzie odpowiednio

$$\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1 \quad \text{lub} \quad \alpha_1 + \alpha_2 + \cdots + \alpha_n + \cdots = 1,$$

to pr-stwo dowolnego zdarzenia A, któremu sprzyja k zdarzeń elementarnych, $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\}$, wyraża się równością:

$$P(A)=\alpha_{i_1}+\alpha_{i_2}+\cdots+\alpha_{i_k}.$$

P10. Jeżeli: 1) $\overline{\overline{\Omega}}=n$ (symbolem $\overline{\overline{Z}}$ oznaczamy liczbę elementów zbioru Z), 2) zdarzenia jednoelementowe są jednakowo prawdopodobne $P(\{\omega_1\})=P(\{\omega_2\})=\ldots=P(\{\omega_n\})=\frac{1}{n}$, 3) zdarzeniu A sprzyja k zdarzeń elementarnych, $\overline{\overline{A}}=k$, to

$$P(A)=\frac{k}{n}=\frac{\text{liczba zdarzeń elementarnych } \textit{sprzyjających} \text{ zdarzeniu A}}{\text{liczba } \textit{wszystkich} \text{ zdarzeń elementarnych przestrzeni } \Omega}$$

Przypomniane w tym twierdzeniu własności pr-stwa, z wyjątkiem własności P8, znane są Czytelnikowi ze szkolnego kursu rachunku pr-stwa. Własność P8 dla n=2 wynika z własności P7; w przypadku ogólnym dowód przeprowadza się metodą indukcji matematycznej.

Własność P10 jest dla nas twierdzeniem dotyczącym wcześniej już określonego pojęcia pr-stwa. Dla matematyka i fizyka francuskiego P.S. Laplace'a własność ta, sformułowana przez niego w 1912r., pełniła rolę definicji pr-stwa. Dlatego własność tę nazywa się **klasyczną definicją pr-stwa** albo **definicją Laplace'a**.

PRZESTRZEŃ PROBABILISTYCZNA. Z dowolnym doświadczeniem związaliśmy już trzy pojęcia: PZE Ω, zbiór zdarzeń $\mathcal{A}$ oraz pr-stwo P określone na tym zbiorze A ; trójkę $(\Omega,\mathcal{A},P)$ nazywa się **przestrzenią probabilistyczną (PP)**; jest ona modelem rozważanego doświadczenia losowego.

PRZYKŁAD 2.1. Pracownik obsługuje trzy tego samego rodzaju maszyny. W ciągu godziny każda z nich co najwyżej raz może wymagać interwencji pracownika z pr-stwem 0,5. Zbudujemy PP $(\Omega,\mathcal{A},P)$ dla tego doświadczenia.

1) **Określamy najpierw PZE** Ω. Jako zdarzenia elementarne ω_i przyjmujemy trzywyrazowe ciągi (i_1,i_2,i_3), gdzie dla $j=1,2,3$, $i_j=1$ albo $i_j=0$ stosownie do tego czy i-ta maszyna wymagała interwencji pracownika czy pracowała bez zarzutu. Zatem

$$\Omega=\{\omega_1=(0,0,0),\ \omega_2=(0,0,1),\ \omega_3=(0,1,0),\ \omega_4=(1,0,0),$$
$$\omega_5=(0,1,1),\ \omega_6=(1,0,1),\ \omega_7=(1,1,0),\ \omega_8=(1,1,1)\}.$$

2) **Wyznaczamy zbiór zdarzeń** $\mathcal{A}$. Zdarzeniami są tu wszystkie podzbiory przestrzeni Ω. Liczba wszystkich zdarzeń dana jest równością $2^8=256$. Mamy tu:

$\binom{8}{0}=1$ - jedno zdarzenie niemożliwe: $A_1=\varnothing$.

$\binom{8}{1}=8$ zdarzeń jednoelementowych:

$A_2=\{\omega_1\}$, $A_3=\{\omega_2\}$, ..., $A_9=\{\omega_8\}$.

$\binom{8}{2}=28$ zdarzeń dwuelementowych postaci $\{\omega_k,\omega_l\}$, gdzie $k,l=1,2,\ldots,8$, $k\neq l$:

$A_{10}=\{\omega_1,\omega_2\}$, $A_{11}=\{\omega_1,\omega_3\}$, ..., $A_{16}=\{\omega_1,\omega_8\}$,

$A_{17}=\{\omega_2,\omega_3\},\ldots,A_{23}=\{\omega_2,\omega_8\},\ldots,A_{37}=\{\omega_7,\omega_8\}$.

$\binom{8}{3}=56$ zdarzeń trzyelementowych postaci $\{\omega_k,\omega_l,\omega_m\}$, gdzie $k,l,m=1,2,\ldots,8$, $k\neq m\neq l$:

$A_{38}=\{\omega_1,\omega_2,\omega_3\}$, ... $A_{93}=\{\omega_6,\omega_7,\omega_8\}$

$\binom{8}{4}=70$ zdarzeń czteroelementowych postaci $\{\omega_k,\omega_l,\omega_m,\omega_s\}$, gdzie $k,l,m,s=1,2,\ldots,8$, $k\neq m\neq l\neq s$:

$A_{94}=\{\omega_1,\omega_2,\omega_3,\omega_4\}$, ... $A_{163}=\{\omega_5,\omega_6,\omega_7,\omega_8\}$.

$\binom{8}{5}=56$ zdarzeń pięcioelementowych:

$A_{164}=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5\}$, ... $A_{219}=\{\omega_4,\omega_5,\omega_6,\omega_7,\omega_8\}$.

$\binom{8}{6}=28$ zdarzeń sześcioelementowych:

$A_{220}=\{\omega_1,\ldots,\omega_6\}$, ... $A_{247}=\{\omega_3,\omega_4,\omega_5,\omega_6,\omega_7,\omega_8\}$.

$\binom{8}{7}=8$ zdarzeń siedmioelementowych:

$A_{248}=\{\omega_1,\ldots,\omega_7\}$, ... $A_{255}=\{\omega_2,\omega_3,\ldots,\omega_8\}$.

$\binom{8}{8}=1$ - jedno zdarzenie pewne: $A_{256}=\Omega=\{\omega_1,\ldots,\omega_8\}$.

3) **Określamy pr-stwo P**, tj. każdemu zdarzeniu A_i przyporządkowujemy pr-stwo $P(A_i)$. Przyjmujemy, że zdarzenia jednoelementowe są jednakowo prawdopodobne: $P(\{\omega_1\})=\ldots=P(\{\omega_8\})=1/8$.

Przy obliczaniu pr-stw $P(A_i)$ korzystamy z własności P9 (lub P10):

$$P(A_1)=P(\emptyset)=0, \qquad P(A_{256})=P(\Omega)=1$$
$$P(A_2)=\ldots=P(A_9)=1/8, \qquad P(A_{10})=\ldots=P(A_{37})=2/8=1/4.$$
$$P(A_{38})=\ldots=P(A_{93})=3/8, \qquad P(A_{94})=\ldots=P(A_{163})=4/8=1/2,$$
$$P(A_{164})=\ldots=P(A_{219})=5/8, \qquad P(A_{220})=\ldots=P(A_{247})=6/8=3/4$$
$$P(A_{248})=\ldots=P(A_{255})=7/8.$$

■

PRAWDOPODOBIEŃSTWO WARUNKOWE. Niech będzie dana PP $(\Omega,\mathcal{A},P)$. Niech $B\in\mathcal{A}$ oraz $P(B)>0$. Dla dowolnego zdarzenia $A\in\mathcal{A}$ liczbę $P(A|B)$ (oznaczaną również symbolem P_B) określoną przy pomocy pr-stwa P równością

$$P(A|B)=\frac{P(A\cap B)}{P(B)}, \qquad A\in\mathcal{A} \tag{2.1}$$

nazywa się **pr-stwem warunkowym zdarzenia** A **przy warunku** B. Funkcję zbioru $P(\cdot|B)$ przyporządkowującą każdemu zdarzeniu $A\in\mathcal{A}$ liczbę $P(A|B)$ zgodnie z (2.1), nazywa się **pr-stwem warunkowym przy warunku** B (w PP $(\Omega,\mathcal{A},P)$).

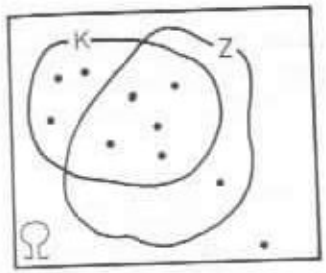
Zauważmy, że pr-stwo P(A) dowolnego zdarzenia A może być traktowane jako pr-stwo warunkowe tego zdarzenia A przy warunku Ω:

$$P(A)=\frac{P(A\cap\Omega)}{P(\Omega)}=P(A|\Omega). \tag{2.2}$$

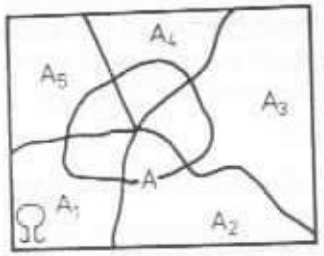
Porównując (2.1) i (2.2) widzimy, że: *pr-stwo warunkowe zdarzenia* A *przy warunku* B *zostało określone tak, jakby zdarzenie* B *pełniło rolę nowej* PZE.

PRZYKŁAD 2.2. Wśród 10-osobowej grupy studenckiej jest ośmiu studentów koleżeńskich, sześciu zdolnych; wiadomo poza tym, że wśród studentów zdolnych jest pięciu koleżeńskich. Doświadczenie polega na losowaniu studenta z tej grupy. Niech K będzie zdarzeniem polegającym na wylosowaniu studenta koleżeńskiego, Z - studenta

zdolnego. Obliczymy P(K|Z) - pr-stwo warunkowe wylosowania studenta koleżeńskiego, jeśli wiadomo, że jest to student zdolny (por. rys 2.1).



Rys 2.1.



Rys 2.2.

Mamy tu PP $(\Omega,\mathcal{A},P)$, gdzie $\Omega=\{$zbiór dziesięciu studentów$\}$, $\mathcal{A}$ - rodzina wszystkich podzbiorów PZE Ω, P - pr-stwo zgodne z definicją klasyczną. Korzystając z (2.1) otrzymujemy:

$$P(K|Z)=\frac{P(K\cap Z)}{P(Z)}=\frac{5/10}{6/10}=\frac{5}{6}.$$

Pr-stwo warunkowe P(K|Z) możemy, zgodnie z wcześniejszym spostrzeżeniem, obliczyć tak jakby zdarzenie Z pełniło rolę nowej PZE i nie odwoływać się do pr-stwa P z PP $(\Omega,\mathcal{A},P)$:

$$P(K|Z)=5/6$$

(jest sześciu studentów zdolnych, $\overline{\overline{Z}}=6$, oraz wśród studentów zdolnych jest pięciu koleżeńskich, $\overline{\overline{K\cap Z}}=5$).

Ten drugi sposób obliczania pr-stwa warunkowego wykorzystuje się przy obliczaniu pr-stwa koniunkcji zdarzeń. ■

PRAWDOPODOBIEŃSTWO KONIUNKCJI. Bezpośrednio ze wzoru (2.1) i przemienności koniunkcji zdarzeń wynika, że *pr-stwo koniunkcji dwóch zdarzeń jest równe iloczynowi pr-stwa dowolnego z tych zdarzeń i pr-stwa warunkowego pozostałego zdarzenia, przy warunku, że to pierwsze zdarzenie zaszło:*

$$P(A\cap B)=\begin{cases}P(A)\cdot P(B|A), & \text{gdy } P(A)>0\\ P(B)\cdot P(A|B), & \text{gdy } P(B)>0\end{cases} \tag{2.3}$$

PRZYKŁAD 2.3. Osoba uchodząca za pechową kupuje dwie sztuki towaru z partii składającej się z siedmiu sztuk dobrych i trzech wadliwych. Wykrycie wad przy zakupie nie jest możliwe. Znajdziemy pr-stwo, że i tym razem nie ominie jej maksymalny pech.

Niech dla $i=1,2$, A_i oznacza zdarzenie - "i-ta zakupiona sztuka jest wadliwa", oraz niech C będzie zdarzeniem - "wszystkie zakupione sztuki są wadliwe". Oczywiście $C=A_1\cap A_2$, więc, zgodnie z wzorem (2.3), mamy:

$$P(C)=P(A_1\cap A_2)=P(A_1)\cdot P(A_2|A_1)=\frac{3}{10}\cdot\frac{2}{9}=\frac{1}{15}.$$ ■

Pr-stwo koniunkcji trzech zdarzeń A, B, C przy założeniu, że $P(A\cap B)>0$ wyraża się następującym wzorem:

(2.4) $$P(A\cap B\cap C)=P(A)\cdot P(B|A)\cdot P(C|A\cap B).$$

Równości (2.3) i (2.4) metodą indukcji uogólnia się na przypadek dowolnej skończonej liczby zdarzeń. Prawdziwe jest mianowicie następujące

TWIERDZENIE 2.2 (o pr-stwie koniunkcji n zdarzeń). Pr-stwo koniunkcji n zdarzeń jest równe iloczynowi pr-stwa jednego z tych zdarzeń przez pr-stwa warunkowe pozostałych zdarzeń obliczone przy warunku, że zaszły wszystkie poprzednie zdarzenia:

(2.5) gdy $P(A_1\cap A_2\cap\ldots\cap A_{n-1})>0$, to $P(A_1\cap A_2\cap\ldots\cap A_n)=$
$=P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1\cap A_2)\cdot\ldots\cdot P(A_n|A_1\cap A_2\cap\ldots\cap A_{n-1})$.

PRZYKŁAD 2.4. Dla danych z przykładu 2.3 znajdziemy pr-stwo, że wśród trzech kupionych sztuk wszystkie będą wadliwe. Zachowując oznaczenia z przykładu otrzymujemy:

$$P(C)=P(A_1\cap A_2\cap A_3)=P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1\cap A_2)=\frac{3}{10}\cdot\frac{2}{9}\cdot\frac{1}{8}=\frac{1}{120}.$$ ■

PRAWDOPODOBIEŃSTWO ZUPEŁNE. Obliczenie pr-stwa zdarzenia A może niekiedy ułatwić następujące

TWIERDZENIE 2.3 (o pr-stwie zupełnym). Jeżeli zdarzenia $A_1,A_2,\ldots,A_n$ tworzą układ zupełny zdarzeń (por. rys 2.2) i mają dodatnie pr-stwa, $P(A_i)>0$, to pr-stwo dowolnego zdarzenia A wyraża się wzorem **(wzorem na pr-stwo zupełne, całkowite)**:

(2.6) $$P(A)=P(A_1)P(A|A_1)+P(A_2)P(A|A_2)+\ldots+P(A_n)P(A|A_n)$$

PRZYKŁAD 2.5. Urządzenie może pracować w dwojakiego rodzaju warunkach: normalnych albo trudnych. Warunki normalne obserwuje się w 70% wszystkich przypadków pracy urządzenia, warunki trudne - 30%. Pr-stwo awarii urządzenia pracującego w warunkach normalnych w czasie T wynosi 0,1; analogicznie pr-stwo dla urządzenia pracującego w warunkach trudnych jest równe 0,4. Obliczymy pr-stwo awarii urządzenia w czasie T.

W tym celu rozważamy zdarzenia:

A_1 - urządzenie będzie pracować w warunkach normalnych,

A_2 - urządzenie będzie pracować w warunkach trudnych,

A - urządzenie ulegnie awarii w czasie T.

Zdarzenia A_1,A_2 spełniają założenia ostatniego twierdzenia, więc na podstawie (2.6), mamy:

$$P(A)=P(A_1)P(A|A_1)+P(A_2)P(A|A_2)=0,7\cdot 0,1+0,3\cdot 0,4=0,19.$$ ■

TWIERDZENIE BAYESA. Pozostając przy założeniach ostatniego twierdzenia załóżmy, że zdarzenie A zaszło. Musiało ono zajść łącznie z jednym spośród zdarzeń $A_1,A_2,\ldots,A_n$. Interesujemy się pr-stwem warunkowym $P(A_k|A)$ tego, że tym zdarzeniem było zdarzenie A_k.

TWIERDZENIE 2.4 (twierdzenie Bayesa). Jeżeli zdarzenia $A_1,A_2,\ldots,A_n$ o dodatnich pr-stwach tworzą układ zupełny zdarzeń oraz zdarzenie A ma również dodatnie pr-stwo, to

(2.7) $$P(A_k|A)=\frac{P(A_k)\cdot P(A|A_k)}{P(A)}\quad\text{dla } k=1,2,\ldots,n,$$

gdzie P(A) dane jest wzorem (2.6).

Dowód. Dwukrotnie stosujemy wzór (2.3) na pr-stwo koniunkcji dwóch zdarzeń:

$$\left.\begin{aligned}P(A_k\cap A)&=P(A_k)\cdot P(A|A_k)\\ P(A_k\cap A)&=P(A)\cdot P(A_k|A)\end{aligned}\right\}\Rightarrow P(A_k)\cdot P(A|A_k)=P(A)\cdot P(A_k|A)\Rightarrow$$

$\Rightarrow$ wzór (2.7). □

P R Z Y K Ł A D 2.6. Dane jak w przykładzie 2.5. Zakładamy, że urządzenie uległo awarii. Wyznaczymy pr-stwo tego, że nastąpiło to przy pracy w trudnych warunkach.

Przy oznaczeniach z przykładu 2.5 należy znaleźć $P(A_2|A)$. Na podstawie twierdzenia Bayesa otrzymujemy:

$$P(A_2|A)=\frac{P(A_2)\cdot P(A|A_2)}{P(A)}=\frac{0{,}3\cdot 0{,}4}{0{,}7\cdot 0{,}1+0{,}3\cdot 0{,}4}=\frac{12}{19}\approx 0{,}63. \quad \blacksquare$$

U w a g a . Występujące w twierdzeniu Bayesa pr-stwa

(a) $$P(A_1),\ P(A_2),\ \ldots,P(A_n)$$

można traktować jako pr-stwa zdarzeń A_i przed rozpoczęciem badania. Dlatego nazywa się je **pr-stwami a priori** zdarzeń $A_1,A_2,\ldots,A_n$ (a priori - znaczy: z góry, przed badaniem).

Pr-stwo zdarzenia jest miarą szansy zajścia tego zdarzenia zależną od posiadanej przez nas informacji. W rezultacie badania zagadnienia wzbogacają się nasze informacje. Mogą zatem zmieniać się pr-stwa zdarzeń związanych z tym zagadnieniem. Tak więc zamiast pr-stw $P(A_1),\ P(A_2),\ \ldots,P(A_n)$ mamy pr-stwa warunkowe

(b) $$P(A_1|A),\ P(A_2|A),\ \ldots,P(A_n|A).$$

Pr-stwa te nazywa się **pr-stwami a posteriori** zdarzeń $A_1,A_2,\ldots,A_n$ (a posteriori - znaczy: z dołu, po badaniu, po doświadczeniu).

NIEZALEŻNOŚĆ ZDARZEŃ. Niech będzie dana PP $(\Omega,\mathcal{A},P)$. Niech $A,B\in\mathcal{A}$. Mówimy, że zdarzenia A, B są **niezależne** (dokładniej: **stochastycznie niezależne**), jeśli

(2.8) $$P(A\cap B)=P(A)\cdot P(B);$$

gdy $P(A\cap B)\neq P(A)\cdot P(B)$, to zdarzenia A, B nazywa się zdarzeniami **zależnymi**.

Z definicji tej i wzoru (2.3) na pr-stwo koniunkcji dwóch zdarzeń, wynika, że:

1) *jeśli* $P(B)>0$, *to*

zdarzenia A, B są niezależne $\Leftrightarrow$ $P(A|B)=P(A)$,

2) *gdy* $P(A)>0$, *to*

zdarzenia A, B są niezależne $\Leftrightarrow$ $P(B|A)=P(B)$.

U w a g a . Nie należy mylić pary zdarzeń n i e z a l e ż n y c h z parą zdarzeń w y k l u c z a j ą c y c h się. Są to pojęcia zupełnie różne chociażby dlatego, że wykluczanie się zdarzeń określa się bez udziału pojęcia pr-stwa (por. zad. 7 c), d)).

Weźmy teraz pod uwagę n zdarzeń $A_1,A_2,\ldots,A_n\in\mathcal{A}$. Mówimy, że zdarzenia te są: 1) **parami niezależne**, gdy każde dwa z nich są niezależne, 2) **zespołowo niezależne** (krócej: **niezależne**), gdy pr-stwo koniunkcji każdych k różnych zdarzeń spośród nich jest równe iloczynowi pr-stw tych zdarzeń.

P R Z Y K Ł A D 2.7. Rozważmy PP $(\Omega,\mathcal{A},P)$, gdzie $\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$, $\mathcal{A}$ - rodzina wszystkich $2^4=16$ podzbiorów przestrzeni Ω, P - pr-stwo wyznaczone warunkiem

$$P(\{\omega_1\})=\ldots=P(\{\omega_4\})=1/4.$$

Rozważmy zdarzenia:

$$A=\{\omega_1,\omega_2\},\quad B=\{\omega_1,\omega_3\},\quad C=\{\omega_1,\omega_4\}.$$

Sprawdzimy, że zdarzenia A, B, C są: a) parami niezależne, b) zespołowo zależne.

a) Ponieważ

$$P(A\cap B)=P(\{\omega_1\})=\frac{1}{4}\ \text{ oraz }\ P(A)\cdot P(B)=\frac{2}{4}\cdot\frac{2}{4}=\frac{1}{4},$$

więc $P(A\cap B)=P(A)\cdot P(B)$ i tym samym zdarzenia A, B są niezależne. Zupełnie analogicznie wykazuje się niezależność zdarzeń A, C oraz niezależność zdarzeń B, C. Zatem zdarzenia A, B, C są parami niezależne.

b) Ponieważ

$$P(A\cap B\cap C)=P(\{\omega_1\})=\frac{1}{4}\ \text{ oraz }\ P(A)\cdot P(B)\cdot P(C)=\frac{2}{4}\cdot\frac{2}{4}\cdot\frac{2}{4}=\frac{1}{8},$$

więc $P(A\cap B\cap C)\neq P(A)\cdot P(B)\cdot P(C)$ i tym samym zdarzenia A, B, C są zespołowo zależne (mimo, że są parami niezależne). $\blacksquare$

ZADANIA DO ROZWIĄZANIA.

1. Wykazać prawdziwość własności P1, P2, ... ,P10 pr-stwa P sformułowanych w twierdzeniu 2.1.
2. Wykazać, że $P(A\cup B)\leq P(A)+P(B)$.
3. Wykazać, że $P[(A\cap B')\cup(A'\cap B)]=P(A)+P(B)-2P(A\cap B)$.
4. Wykazać, że $P(A\cap B')-P(A'\cap B)=P(A)-P(B)$.

5. Czy zdarzenia A i B mogą być rozłączne, jeśli:
   a) $P(A)=0{,}33$, $P(B)=0{,}72$, b) $P(A)=0{,}49$, $P(B)=0{,}48$,
   c) $P(A)>0{,}49$, $P(B)>0{,}49$, d) $P(A)>0{,}33$, $P(B)\geq 0{,}67$.
6. Dla jakich wartości parametru t zdarzenia A, B mogą być rozłączne, jeśli $P(A)=4t$, $P(B)=5t^2$.
7. Wykazać, że
   a) dowolne zdarzenie A i zdarzenie B o pr-stwie $P(B)=0$ są niezależne, b) dowolne zdarzenie A i zdarzenie B o pr-stwie $P(B)=1$ są niezależne,
   c) dwa rozłączne zdarzenia o dodatnich pr-stwach są zależne,
   d) dwa rozłączne zdarzenia, z których przynajmniej jedno ma zerowe pr-stwo, są niezależne,
   e) jeśli zdarzenia A, B są niezależne, to niezależne są również pary zdarzeń: A' i B, A i B', A' i B',
   f) jeśli $P(A|B)=P(A|B')$, to zdarzenia A, B są niezależne.
8. Zdarzenia A, B są niezależne oraz $P(A)=0{,}2$, $P(B)=0{,}5$. Obliczyć pr-stwa zajścia: a) obydwu zdarzeń A i B, b) co najmniej jednego z tych zdarzeń, c) dokładnie jednego z tych zdarzeń.
9. Zdarzenia A, B są niezależne oraz $P(A)=0{,}2$, $P(B)=0{,}5$. Obliczyć pr-stwa: a) $P(A\cap B)$, b) $P(A\cup B)$, c) $P(A\setminus B)$, d) $P(A|B)$, e) $P(A|A\cap B)$, f) $P(A|A\cup B)$.
10. Obliczyć pr-stwa z zadania 9 przy założeniu, że zdarzenia A, B są rozłączne.
11. Lokalizację celu przeprowadzają niezależnie od siebie dwie stacje radiolokacyjne. Pr-stwa lokalizacji celu przez te stacje są odpowiednio równe 0,8 i 0,9. Obliczyć pr-stwo, że cel zostanie zlokalizowany co najmniej przez jedną stację.
12. Zdarzenia A, B są niezależne oraz ich alternatywa jest zdarzeniem pewnym. Wykazać, że wtedy $P(A)=1$ lub $P(B)=1$.
13. Urządzenie przestaje działać, gdy uszkodzi się jeden z jego n jednorodnych elementów. Usuwanie awarii odbywa się w ten sposób, że uszkodzony element wymienia się na nowy. Zakładając, że pr-stwo uszkodzenia poszczególnych elementów jest jednakowe i wynosi p, obliczyć pr-stwo, że trzeba będzie sprawdzać wszystkie n elementów.

14. Dziecko bawi się klockami z literami tworzącymi słowo WADA, układając trzy z nich obok siebie. Obliczyć pr-stwo ułożenia w ten sposób słowa DWA.
15. Kandydat na studia ma zdawać egzamin wstępny z matematyki i fizyki według następującej umowy: otrzyma 3 pytania na przemian z tych przedmiotów i zostanie przyjęty, jeśli odpowie poprawnie na dwa kolejne pytania. Od jakiego przedmiotu powinien rozpocząć egzamin, jeśli matematykę ma opanowaną w 90% a fizykę w 60%?.
16. W urnie jest n losów w tym k wygrywających. Z urny tej jeden los wyjmuje pierwsza osoba, a następnie z pozostałych losów również jeden los wyjmuje druga osoba. Sprawdzić, że pr-stwo wylosowania wygrywającego losu jest jednakowe dla obu osób.
17. W każdej z dwu partii tego samego rodzaju towaru, liczących odpowiednio m i n sztuk, znajduje się po jednej sztuce wadliwej. Losujemy jedną sztukę z pierwszej partii i przekładamy (bez oglądania) do drugiej. Następnie z drugiej partii losujemy jedną sztukę. Obliczyć pr-stwo, że jest to sztuka dobra.
18. Dwa automaty produkują jednakowe detale, które dostarczane są na przenośnik taśmowy. Produkcja pierwszego automatu jest dwa razy większa niż produkcja drugiego. Pierwszy automat produkuje średnio 60% detali pierwszej jakości, a drugi 84% detali pierwszej jakości. Wyznaczyć pr-stwo tego, że losowo wzięty detal, który okazał się pierwszej jakości, został wyprodukowany przez pierwszy automat.
19. Trzy ściany czworościanu są pomalowane odpowiednio na czerwono, niebiesko i zielono, czwarta - na wszystkie trzy kolory. Niech C, N, Z oznaczają zdarzenia: przy rzucaniu czworościanu na płaszczyznę czworościan upadł ścianą zawierającą kolor czerwony, niebieski, zielony. Sprawdzić, że zdarzenie C, N, Z są parami niezależne ale nie są niezależne zespołowo.

Odpowiedzi.

5. a) nie, b) tak, c) tak (gdy $P(A)+P(B)\leq 1$), d) nie. 6. $0\leq t\leq 0{,}2$.
8. a) 0,1, b) 0,6, c) 0,5. 9. a) 0,1, b) 0,6, c) 0,1, d) 0,2, e) 1, f) 1/3.
10. a) 0, b) 0,7, c) 0,2, d) 0, e) nie istnieje, f) 2/7. 11. 0,98. 13. $(1-p)^{n-1}$.
14. 1/12. 15. Rozpocząć od pytania z fizyki. 17. $(mn-1)/m(n+1)$. 18. 10/17.

3. ZMIENNE LOSOWE.

POJĘCIE ZMIENNEJ LOSOWEJ. W przykładach i zadaniach widzieliśmy, że zdarzenia elementarne ω mogą mieć bardzo różnorodną naturę. W celu uwolnienia się od tej różnorodności "zastępuje się" je liczbami. Realizuje się to drogą wprowadzenia zmiennej losowej.

Zmienną losową (ZL) nazywa się funkcję o wartościach ze zbioru R liczb rzeczywistych, której dziedziną jest PZE Ω. Innymi słowy ZL, to odwzorowanie PZE Ω w zbiór R liczb rzeczywistych.

ZL będziemy oznaczać dużymi literami X, Y, Z, a ich **wartości** (nazywane również **realizacjami**) X(w), Y(w), Z(w) odpowiednio małymi literami x, y, z.

Dodajmy, że w pełnym wykładzie rachunku pr-stwa na funkcje, które nazwaliśmy zmiennymi losowymi, nakłada się pewien warunek, który w rozważanych przez nas zagadnieniach jest zawsze spełniony.

PRZYKŁAD 3.1 (c.d. przykładu 2.1). Pracownik obsługuje trzy maszyny. W ciągu godziny każda z nich co najwyżej raz może wymagać interwencji pracownika. Niech ZL X oznacza liczbę maszyn, które w ciągu godziny wymagają interwencji pracownika. Mamy tu (zachowujemy oznaczenie z przykładu 2.1):

$$\Omega=\{\omega_1=(0,0,0),\ \omega_2=(0,0,1),\ \omega_3=(0,1,0),\ \omega_4=(1,0,0),$$
$$\omega_5=(0,1,1),\ \omega_6=(1,0,1),\ \omega_7=(1,1,0),\ \omega_8=(1,1,1)\}.$$

ZL X określona jest następująco:

$$X(\omega_1)=0,\quad X(\omega_2)=X(\omega_3)=X(\omega_4)=1,$$
$$X(\omega_5)=X(\omega_6)=X(\omega_7)=2,\quad X(\omega_8)=3.$$

Tak więc ZL X przekształca PZE Ω na zbiór liczbowy $W=\{0,1,2,3\}$. ■

ROZKŁAD PRAWDOPODOBIEŃSTWA ZMIENNEJ LOSOWEJ. Niech będzie dana PP $(\Omega,\mathcal{A},P)$. Jeżeli X jest ZL, zaś A zbiorem na prostej, to relację $X\in A$ odczytujemy : "ZL X przyjmie wartość ze zbioru A" (a nie "X należy do A", bo przecież A nie jest zbiorem zmiennych losowych). W szczególności relacje: $a<X<b$, $X<b$, $X=x_0$ odczytujemy odpowiednio: ZL X przyjmie wartość z przedziału (a,b), mniejszą od b, równą liczbie x_0.

Z tymi relacjami chcemy związać pr-stwo P ich realizacji (poprawniej byłoby oznaczyć to pr-stwo innym symbolem, np. P_X).

Przyjmujemy:

$$P(X\in A)\overset{\text{def}}{=}P\{\omega:\ X(\omega)\in A\},\tag{3.1}$$

czyli: ***pr-stwo*** $P(X\in A)$ ***przyjęcia przez*** ZL X ***wartości ze zbioru*** A, $A\subset R$, *określamy jako pr-stwo zdarzenia* $\{\omega:\ X(\omega)\in A\}$, *składającego się z tych zdarzeń elementarnych* ω, *dla których* ZL X *przyjmuje wartości* $X(\omega)$ *ze zbioru* A.

Jeżeli dla każdego zbioru $A\subset R$ dane jest pr-stwo $P(X\in A)$, to mówimy, że dany jest **rozkład pr-stwa ZL** X.

PRZYKŁAD 3.2. W przykładzie 2.1 zbudowaliśmy PP $(\Omega,\mathcal{A},P)$ dla doświadczenia polegającego na obsłudze trzech maszyn. Z kolei w ostatnim przykładzie określiliśmy ZL X związaną z tym doświadczeniem. Obecnie obliczymy pr-stwo przyjęcia przez ZL X poszczególnych wartości 0, 1, 2, 3 oraz pr-stwo tego, że "co najmniej dwie maszyny będą wymagać interwencji pracownika", czyli $P(X\geq 2)$. Zgodnie z definicją (3.1), mamy:

$$P(X=0)=P(\{\omega_1\})=1/8,\quad P(X=1)=P(\{\omega_2,\omega_3,\omega_4\})=3/8,$$
$$P(X=2)=P(\{\omega_5,\omega_6,\omega_7\})=3/8,\quad P(X=3)=P(\{\omega_8\})=1/8,$$
$$P(X\geq 2)=P(\{\omega_5,\omega_6,\omega_7,\omega_8\})=4/8=1/2.$$ ■

ZMIENNE LOSOWE TYPU SKOKOWEGO. Wyróżnia się dwa ważne typy zmiennych losowych: zmienne losowe typu skokowego (dyskretnego, ziarnistego) i zmienne typu ciągłego. Odpowiadające im rozkłady pr-stwa nazywa się również odpowiednio: rozkładami pr-stwa typu skokowego i rozkładami pr-stwa typu ciągłego. Dodajmy, że te dwa typy nie wyczerpują wszystkich możliwych typów zmiennych losowych i ich rozkładów pr-stwa.

Wartość x_i ZL X nazywa się **punktem skokowym** (lub **atomem**) ZL X, jeśli pr-stwo $P(X=x_i)$, oznaczane również przez p_i lub $p(x_i)$, jest dodatnie:

$$P(X=x_i)\equiv p(x_i)\equiv p_i>0.$$

Przy tym liczbę p_i nazywa się **skokiem** ZL X w punkcie skokowym x_i. Punkty skokowe x_i i skoki p_i ZL X nazywa się również **punktami skokowymi** i **skokami rozkładu pr-stwa ZL X**.

Dowodzi się, że *zbiór* W *punktów skokowych* x_i *może być skończony albo przeliczalny*. Zatem zbiór W punktów skokowych jest postaci:

$$W=\{x_1,x_2,\ldots,x_n\} \quad \text{lub} \quad W=\{x_1,x_2,\ldots,x_n,\ldots\}.$$

Mówimy, że X jest **zmienną losową skokową (ZLS)** jeśli: 1) zbiór W jej punktów skokowych x_i jest skończony albo przeliczalny oraz 2) suma wszystkich jej skoków p_i spełnia warunek unormowania, tj.

$$\sum_{x_i \in W} p_i = 1. \tag{3.2}$$

Jeśli punkty skokowe x_i ZLS X dają się ustawić w ciąg ściśle monotoniczny, to *między każdymi dwoma kolejnymi punktami skokowymi istnieje cały przedział bez punktów skokowych*. Jest zrozumiałe, że ZLS X poza punktami skokowymi może przyjmować inne wartości x ale z pr-stwem zero: $P(X=x)=0$.

Jeśli X jest ZLS, to funkcję $p(\cdot)$ określoną na zbiorze W jej punktów skokowych x_i równością

$$p(x_i) \overset{def}{=} P(X=x_i)\equiv p_i \tag{3.3}$$

lub tabelą

(3.3')

x_i	x_1	x_2	....	x_n	....
p_i	p_1	p_2	....	p_n	....

nazywa się **funkcją pr-stwa ZLS X**.

Funkcję pr-stwa ZLS X można interpretować w fizyce, jako rozmieszczenie (rozkład) masy jednostkowej w pojedynczych, izolowanych punktach x_i osi 0x, przy czym masa $m_i=p(x_i)=p_i$ zostaje skoncentrowana w punkcie o odciętej x_i.

Wykresem funkcji pr-stwa ZLS X jest zbiór punktów (x_i,p_i), gdzie $x_i \in W$. Jeśli punkty te połączyć odcinkami z ich rzutami na oś 0x, to otrzymuje się figurę nazywaną niekiedy **histogramem funkcji pr-stwa**.

Gdy $A \subset R$, to na mocy definicji przyjmujemy:

$$P(X \in A) \overset{def}{=} \sum_{x_i \in A} p_i \tag{3.4}$$

(pr-stwo przyjęcia przez ZLS X wartości ze zbioru A jest równe sumie skoków p_i, odpowiadających punktom skokowym x_i, które "trafiły" do zbioru A). Widać stąd, że *funkcja pr-stwa* ZLS X *za pomocą wzoru* (3.4) *w pełni wyznacza rozkład pr-stwa tej* ZL.

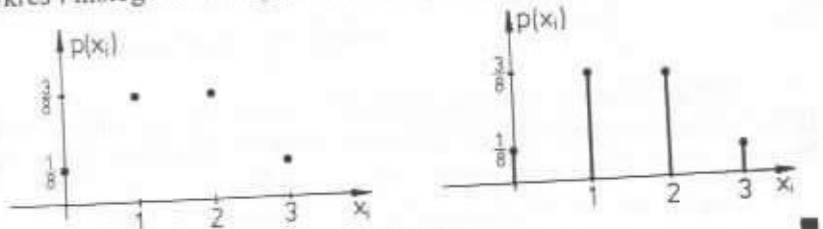
PRZYKŁAD 3.3. ZL X z przykładu 3.1 jest ZLS a jej funkcja pr-stwa jest postaci:

x_i	0	1	2	3
p_i	1/8	3/8	3/8	1/8

Pr-stwo $P(X\le 2)$ wyraża się równością postaci (3.4):

$$P(X\le 2)=\sum_{x_i\le 2} p_i=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}=\frac{7}{8}.$$

Wykres i histogram funkcji pr-stwa znajdują się na rysunku 3.1.



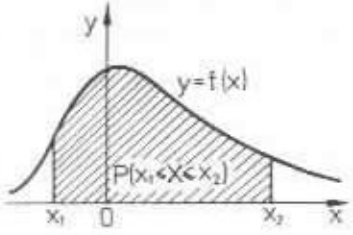
Rys 3.1.

ZMIENNE LOSOWE TYPU CIĄGŁEGO. Mówimy, że X jest **zmienną losową typu** ciągłego albo krócej **zmienną losową ciągłą (ZLC)**, jeśli istnieje nieujemna i całkowalna na całej osi funkcja f taka, że dla każdego przedziału $<x_1,x_2>$

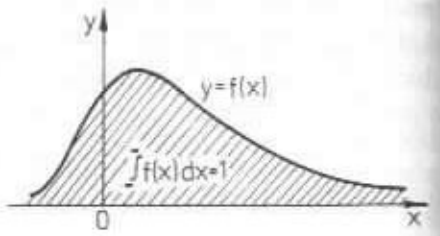
$$P(x_1 \le X \le x_2)=\int_{x_1}^{x_2} f(x)dx\,; \tag{3.5}$$

przy tym funkcję f nazywa się **gęstością prawdopodobieństwa (GP) ZL X**, a jej wykres - **krzywą gęstości ZL X**.

Z (3.5), gdy uwzględnić interpretację geometryczną całki oznaczonej wynika, że *pr-stwo przyjęcia przez* ZLC *wartości z pewnego przedziału utożsamiamy (liczbowo) z polem pod krzywą gęstości nad tym przedziałem* (por. rys 3.2).



Rys 3.2.



Rys 3.3.

GP ZLC można interpretować w fizyce jako *gęstość liniową masy jednostkowej* rozmieszczonej (chciałoby się powiedzieć "rozsmarowanej") wzdłuż osi 0x.

Rozważmy przypadki szczególne równości (3.5) przyjmując najpierw $x_1=-\infty$, $x_2=+\infty$. Ponieważ

$$P(-\infty<X<+\infty)=P(\{\omega:-\infty<X(\omega)<+\infty\})=P(\Omega)=1, \quad \text{więc}$$

$$\int_{-\infty}^{\infty} f(x)dx=1 \qquad \text{(warunek unormowania).} \tag{3.6}$$

GP ZLC jest to zatem *nieujemna i całkowalna na całej osi funkcja* f *spełniająca warunek unormowania* (3.6). *Pole nad osią* 0x *pod krzywą gęstości jest równe jedności* (por. rys 3.3).

Niech teraz w równości (3.5) $x_1=x_2$, wtedy przedział $<x_1,x_2>$ redukuje się do zbioru jednopunktowego $<x_1,x_2>=\{x_1\}$ i lewa jej strona jest pr-stwem $P(X=x_1)$, prawa zaś wynosi zero. Zatem

$$P(X=x_1)=0, \tag{3.7}$$

czyli: *pr-stwo przyjęcia przez ZLC pojedynczej, ustalonej wartości jest zawsze równe zero (ZLC nie ma punktów skokowych!).* Dlatego dla ZLC celowe staje się dopiero pytanie o pr-stwo tego, że przyjmie ona wartości z (niezdegenerowanego) przedziału; przy tym

$$P(x_1\le X\le x_2)=P(x_1\le X<x_2)=P(x_1<X\le x_2)=P(x_1<X<x_2).$$

Wróćmy jeszcze do równości (3.7). Początkowo może wydawać się zaskakujące, że mimo iż poszczególne wartości x przyjmowane są przez ZLC z zerowym pr-stwem, to jednak można z nich "złożyć" przedział, powiedzmy $<x_1,x_2>$, $x_1<x_2$, taki, że pr-stwo $P(x_1\le X\le x_2)$ jest dodatnie. Okoliczność ta przestanie nas dziwić, gdy przypomnimy sobie, że np. mimo, iż poszczególne punkty jakiegoś przedziału, powiedzmy $<0,5>$, są zerowej długości, to jednak punkty te tworzą przedział o dodatniej długości 5.

PRZYKŁAD 3.4. GP ZLC X jest postaci

$$f(x)=\begin{cases}A(2x-x^2) & \text{dla } 0<x<2\\ 0 & \text{dla pozostałych } x\end{cases}$$

(o takiej GP mówi się, że jest **skoncentrowana** na przedziale (0,2) albo, że przedział (0,2) jest **nośnikiem** rozkładu ZL X). Wyznaczymy:

a) stałą A, b) $P(0\le X\le 1/2)$, c) $P(0\le X\le 1)$.

a) stałą A znajdujemy z warunku unormowania (3.6):

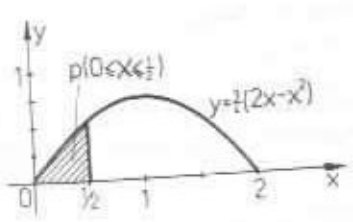
$$\int_{-\infty}^{0}0dx+\int_{0}^{2}A(2x-x^2)dx+\int_{2}^{\infty}0dx=1.$$

Stąd $4A/3=1$, czyli $A=3/4$.

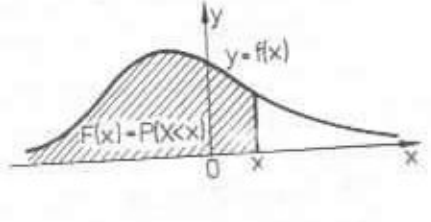
b) Na podstawie równości (3.5) znajdujemy (por. rys 3.4),

$$P(0\le X\le\tfrac{1}{2})=\int_{0}^{1/2}\frac{3}{4}(2x-x^2)dx=\frac{3}{4}(x^2-\frac{1}{3}x^3)\Big|_0^{1/2}=\frac{5}{32}.$$

c) Krzywa gęstości jest symetryczna względem prostej $x=1$. Stąd oraz z warunku unormowania znajdujemy bez całkowania szukane pr-stwo: $P(0\le X\le 1)=1/2$. ■



Rys 3.4.



Rys 3.5.

DYSTRYBUANTA. Niech X będzie ZL dowolnego typu określoną na PP $(\Omega,\mathcal{A},P)$.

Dystrybuantą ZL X nazywamy funkcję F określoną na całej osi równością:

$$F(x)\overset{def}{=}P(X<x), \quad x\in R. \tag{3.8}$$

Jeśli X jest ZLS o punktach skokowych x_k i skokach $p_k = p(x_k)$, to wzór (3.8) przyjmuje następującą postać:

(3.9) $$F(x)=\sum_{x_k<x} p(x_k).$$

W szczególności, gdy $x_1, x_2, \ldots, x_n$ są kolejnymi punktami skokowymi, to wzór (3.9) można zapisać w równoważnej postaci:

(3.9)' $$F(x)=\begin{cases}0 & \text{dla } x\le x_1\\ F(x_k)+p(x_k) & \text{dla } x\in(x_k,x_{k+1}>,\ k=1,2,\ldots,n-1\\ 1 & \text{dla } x>x_n\end{cases}$$

Jeśli X jest ZLC o GP f, to wzór (3.8) przyjmuje następującą postać, (por. rys.3.5),

(3.10) $$F(x)=\int_{-\infty}^{x} f(t)dt,\quad x\in R.$$

Korzystając z definicji dystrybuanty dowodzi się

TWIERDZENIE 3.1. Dystrybuanta F dowolnej ZL X ma następujące własności F1, F2,...,F6:

F1. $0\le F(x)\le 1$ dla każdego $x\in R$.

F2. Dystrybuanta F jest funkcją niemalejącą.

F3. $\lim_{x\to-\infty} F(x)\equiv F(-\infty)=0$, $\lim_{x\to+\infty} F(x)\equiv F(+\infty)=1$.

F4. Dystrybuanta F jest funkcją lewostronnie ciągłą, czyli $F(x_0^-)=F(x_0)$, gdzie $F(x_0^-)$ oznacza lewostronną granicę dystrybuanty F w punkcie x_0.

F5. $F(x_2)-F(x_1)=P(x_1\le X<x_2)$ dla $x_1<x_2$.

F6. $F(x_0^+)-F(x_0)=P(X=x_0)$ dla dowolnego $x_0\in R$.

Własności F1-F6 to warunki konieczne na to, by funkcja F mogła być dystrybuantą. Okazuje się, że warunki F2, F3, F4 również wystarczają na to, aby z góry zadana funkcja F była dystrybuantą. Prawdziwe jest bowiem

TWIERDZENIE 3.2. Jeżeli funkcja F określona na całej osi ma własności F2, F3, F4 sformułowane w twierdzeniu 3.1, to jest ona dystrybuantą.

Własności F1-F6 dotyczą ZL dowolnego typu. Załóżmy teraz, że F jest dystrybuantą ZLC X. Można oczekiwać, że własności te mogą przybrać teraz bardziej szczegółową postać lub mogą pojawiać się nowe. Traktuje o tym następujące

TWIERDZENIE 3.3. Dystrybuanta F ZLC ma następujące własności FC1, FC2, FC3:

FC1. Gdy $x_1<x_2$, to

$$F(x_2)-F(x_1)=P(x_1\le X<x_2)=P(x_1<X\le x_2)=P(x_1\le X\le x_2)=P(x_1<X<x_2).$$

FC2. Dystrybuanta ZLC jest funkcją ciągłą na całej osi.

FC3. Jeśli F jest dystrybuantą ZLC X, to funkcja f postaci

(3.11) $$f(x)=\begin{cases}F'(x) & \text{dla tych } x, \text{ dla których istnieje pochodna } F'(x)\\ 0 & \text{dla pozostałych } x\end{cases}$$

jest GP ZL X.

Własność FC3 z twierdzenia 3.3 umożliwia wyznaczanie GP f, gdy wiemy, że F jest dystrybuantą ZLC. Z następnego twierdzenia wynika sposób rozpoznawania dystrybuanty ZLC.

TWIERDZENIE 3.4. Jeżeli F jest dystrybuantą ZL X i funkcja f określona równością postaci (3.11) spełnia warunek unormowania (3.6), to X jest ZLC oraz funkcja f jest jej GP.

PRZYKŁAD 3.5. Wyznaczymy dystrybuantę ZLS X z przykładu 3.3. Jej funkcja pr-stwa jest postaci:

x_k	0	1	2	3
$p_k=p(x_k)$	1/8	3/8	3/8	1/8

Wykorzystamy wzór (3.9)'. Ponieważ ZLS X ma cztery punkty skokowe, więc jej dystrybuantę należy rozpatrywać w pięciu przedziałach, na jakie dzielą one oś 0x. Kolejno obliczamy:

dla $x\le 0$ $F(x)=P(X<x)=P(X<0)=F(0)=0$,

dla $0<x\le 1$

$$F(x)=P(X<x)=P(X<1)=F(1)=F(0)+p(0)=0+\frac{1}{8}=\frac{1}{8},$$

dla $1<x\le 2$

$$F(x)=P(X<x)=P(X<2)=F(2)=F(1)+p(1)=\frac{1}{8}+\frac{3}{8}=\frac{4}{8},$$

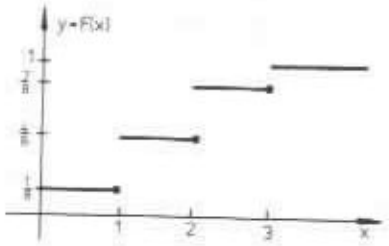
dla $2<x\le 3$

$$F(x)=P(X<x)=P(X<3)=F(3)=F(2)+p(2)=\frac{4}{8}+\frac{3}{8}=\frac{7}{8},$$

dla $x>3$ $\quad F(x)=P(X<x)=P(X<3+\varepsilon)=F(3)+p(3)=\frac{7}{8}+\frac{1}{8}=1.$

"Zbierając" wyniki otrzymujemy poszukiwaną dystrybuantę F (rys 3.6):

x	$(-\infty,0>$	$(0,1>$	$(1,2>$	$(2,3>$	$(3,\infty)$
F(x)	0	1/8	4/8	7/8	1



Rys 3.6. ■

Widzimy, że dystrybuanta ZLS X o punktach skokowych $x_1<x_2<\ldots<x_n$ jest **funkcją schodkową**: jest stała w przedziałach $(-\infty,x_1>,(x_1,x_2>,\ldots,(x_{n-1},x_n>,(x_n,+\infty)$ i ma skoki p_k w tych punktach skokowych x_k; wewnątrz tych przedziałów jest funkcją ciągłą, w punktach skokowych x_k jest lewostronnie ciągła i nie jest prawostronnie ciągła.

P R Z Y K Ł A D 3.6. Wyznaczymy teraz dystrybuantę ZLC X z przykładu 3.4. Jej GP jest funkcja

$$f(x)=\begin{cases}\frac{3}{4}(2x-x^2) & \text{dla } 0<x<2\\ 0 & \text{dla } x\le 0 \vee x\ge 2.\end{cases}$$

Wykorzystujemy wzór (3.10) wyznaczający dystrybuantę ZLC. Ponieważ GP jest skoncentrowana na przedziale (0,2), więc

$F(x)=0$ dla $x\le 0$, $\quad F(x)=1$ dla $x>2$

oraz dla $x\in(0,2]$

$$F(x)=\int_{-\infty}^{x}f(t)dt=\int_{-\infty}^{0}0dt+\int_{0}^{x}\frac{3}{4}(2t-t^2)dt=\frac{1}{4}(3x^2-x^3).$$

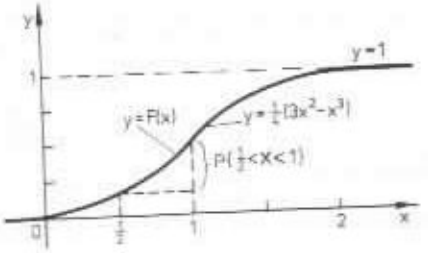
Zbierając te wyniki, otrzymujemy dystrybuantę F ZLC X:

$$F(x)=\begin{cases}0 & \text{dla } x\le 0\\ \frac{1}{4}(3x^2-x^3) & \text{dla } 0<x\le 2\\ 1 & \text{dla } x>2\end{cases}$$

lub w postaci tabeli

x	$(-\infty,0>$	$(0,2>$	$(2,\infty)$
F(x)	0	$(3x^2-x^3)/4$	1

Wykres wyznaczonej dystrybuanty jest pokazany na rysunku 3.7.



Rys 3.7.

Korzystając bezpośrednio z wyznaczonej dystrybuanty obliczymy jeszcze pr-stwo $P(1/2<X<1)$. Szukane pr-stwo jest przyrostem dystrybuanty między punktami 1/2 i 1, (por. rys.3.7):

$$P(\frac{1}{2}<X<1)=F(1)-F(\frac{1}{2})=\frac{1}{4}[(3-1)-(\frac{3}{4}-\frac{1}{8})]=\frac{11}{32}.$$ ■

W ostatnich dwu przykładach na podstawie danej funkcji pr-stwa ZLS i danej GP ZLC wyznaczyliśmy dystrybuantę F. W dalszych dwu przykładach postąpimy odwrotnie: mając daną dystrybuantę ZL wyznaczymy jej funkcję pr-stwa, gdy jest to ZLS albo jej GP, gdy jest to ZLC.

P R Z Y K Ł A D 3.7. Dana jest dystrybuanta ZLS X:

x	$(-\infty,-2>$	$(-2,0>$	$(0,2>$	$(2,3>$	$(3,\infty)$
F(x)	0	0,2	0,4	0,8	1

Wyznaczymy jej funkcję pr-stwa.

Punktami skokowymi x_i ZL X są punkty nieciągłości dystrybuanty, tj punkty -2, 0, 2, 3. Skoki $p_i = p(x_i) = P(X = x_i)$ znajdujemy na podstawie własności F6 dystrybuanty. Na przykład:

$$p_2 = p(0) = P(X=0) = F(0^+) - F(0) = 0{,}4 - 0{,}2 = 0{,}2.$$

Analogicznie otrzymujemy:

$p_1 = p(0) = P(X=-2) = 0{,}2$, $p_3 = P(X=2) = 0{,}4$, $p_4 = P(X=3) = 0{,}2$.

Jak zwykle wyznaczoną funkcję pr-stwa zapisujemy w postaci tabeli:

x_i	-2	0	2	3
p_i	0,2	0,2	0,4	0,2

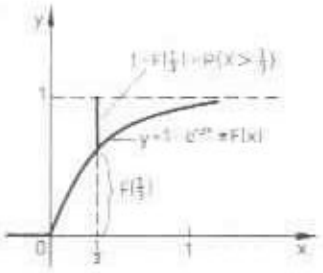
■

PRZYKŁAD 3.8. Dana jest dystrybuanta F ZL X, (por. rys.3.8),

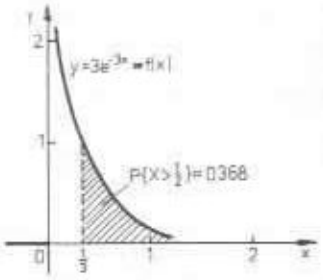
$$F(x) = \begin{cases} 0 & \text{dla } x \le 0 \\ 1 - e^{-3x} & \text{dla } x > 0. \end{cases}$$

a) Sprawdzimy, że X jest ZLC i wyznaczymy jej GP f.

b) Obliczymy pr-stwo $P(X > \frac{1}{3})$ korzystając tylko z dystrybuanty i podamy jego interpretację na wykresie dystrybuanty i GP f.



Rys 3.8



Rys 3.9

a) Dana dystrybuanta jest funkcją różniczkowalną na całej osi z wyjątkiem punktu $x = 0$:

$F'(x) = 0$ dla $x < 0$ oraz $F'(x) = 3e^{-3x}$ dla $x > 0$.

Funkcja

$$f(x) = \begin{cases} 0 & \text{dla } x \le 0 \\ 3e^{-3x} & \text{dla } x > 0 \end{cases}$$

jest postaci (3.11) i spełnia warunek unormowania (3.6). Zatem zgodnie z twierdzeniem 3.4, X jest ZLC o GP f, (por. rys.3.9).

b) Suma pr-stw zdarzeń przeciwnych jest równa jedności, zatem

$$P(X > \tfrac{1}{3}) = 1 - P(X \le \tfrac{1}{3}) = 1 - P(X < \tfrac{1}{3}) = 1 - F(\tfrac{1}{3}) = e^{-1} \approx 0{,}368.$$ ■

PRZYKŁAD 3.9. Dobierzemy tak stałe a i b aby funkcja F dana wzorem

$$F(x) = \begin{cases} 0 & \text{dla } x \le 0 \\ ax + b & \text{dla } 0 < x \le 1 \\ 1 & \text{dla } x > 1 \end{cases}$$

była dystrybuantą: a) ZLC, b) ZLS.

a) Funkcja F jako dystrybuanta ZLC jest ciągła (obustronnie) na całej osi, a więc również w punktach $x = 0$ i $x = 1$:

$$\lim_{x \to 0+} F(x) = F(0) \Leftrightarrow \lim_{x \to 0+} (ax + b) = F(0) \Leftrightarrow a \cdot 0 + b = 0 \Leftrightarrow b = 0,$$

$$\lim_{x \to 1+} F(x) = F(1) \Leftrightarrow \lim_{x \to 1+} 1 = F(1) \Leftrightarrow 1 = a \cdot 1 + b \Leftrightarrow a = 1.$$

Jeśli zatem F ma być dystrybuantą ZLC, to należy przyjąć $a = 1$, $b = 0$.

b) Funkcja F jako dystrybuanta ZLS musi być przedziałami stała. Dlatego w tym przypadku należy przyjąć $a = 0$. Jest też zrozumiałe, że b może być dowolną liczbą z przedziału $<0,1>$. Zatem, jeśli $a = 0$, $b \in <0,1>$, to dana funkcja F jest dystrybuantą ZLS. ■

ZADANIA DO ROZWIĄZANIA.

1. Dana jest funkcja pr-stwa ZL X:

x_i	1	3	5	7	10
p_i	0,1	0,1	c	0,2	0,3

Wyznaczyć: a) stałą c, b) wykres i histogram funkcji pr-stwa, c) dystrybuantę i jej wykres, d) pr-stwa: $p_1 = P(X=3)$, $p_2 = P(X=2)$, $p_3 = P(X<7)$, $p_4 = P(X \le 10)$, $p_5 = P(X>3)$, $p_6 = P(3 \le X < 7)$, $p_7 = P(-5 \le X < 7)$, $p_8 = P(|X-5|<3)$, $p_9 = P(|X-3| \ge 1)$ dwiema metodami korzystając: 1) tylko z danej funkcji pr-stwa oraz 2) z wyznaczonej dystrybuanty; znalezione pr-stwa p_i zaznaczyć na rysunku z wykresem dystrybuanty.

2. Dana jest dystrybuanta ZLS X:

x	$(-\infty,-3>$	$(-3,0>$	$(0,4>$	$(4,9>$	$(9,15>$	$(15,\infty)$
F(x)	0	0,2	0,4	0,7	0,8	1

Wyznaczyć funkcję pr-stwa ZL X.

3. a) Wyrazić za pomocą dystrybuanty następujące pr-stwa:
$p_1 = P(X \geq b)$, $p_2 = P(X \leq b)$, $p_3 = P(a < X \leq b)$, $p_4 = P(a \leq X \leq b)$, $p_5 = P(a < X < b)$, gdy X jest ZLS;
b) uprościć odpowiedzi do zadania a), gdy X jest ZLC;
c) korzystając z a) obliczyć pr-stwa: $p_1 = P(X \leq 9)$, $p_2 = P(4 \leq X \leq 15)$, $p_3 = P(4 < X < 15)$, gdzie X jest ZL z zadania 2.

4. Niech w warunkach przykładu 3.1 Y oznacza liczbę maszyn, które w ciągu godziny nie wymagały interwencji pracownika. Wyznaczyć funkcję pr-stwa ZL Y. Porównać otrzymany wynik z funkcją pr-stwa ZL X wyznaczoną w przykładzie 3.3. Uzupełnić zdanie: "dwie różne ZLS mogą mieć jednakowe ...".

5. Ze zbioru siedmiu odcinków, wśród których trzy są długości 3 cm każdy i cztery mają długość 5 cm każdy, losujemy trzy odcinki. Każdemu trójkątowi zbudowanemu z tych odcinków przyporządkujemy obwód. Wyznaczyć: a) funkcję pr-stwa, b) dystrybuantę tak określonej ZL.

6. Spośród wierzchołków sześciokąta foremnego o boku długości 1 cm wybrano losowo trzy różne. Każdemu trójkątowi o wierzchołkach w wylosowanych punktach przyporządkowujemy jego pole. Wyznaczyć: a) funkcję pr-stwa, b) dystrybuantę tak określonej ZL.

7. Dana jest funkcja f: $f(x) = \begin{cases} a(1-|x-1|) & \text{dla } 0 < x < 2 \\ 0 & \text{dla } x \leq 0 \vee x \geq 2 \end{cases}$

a) Wyznaczyć stałą a tak, aby funkcja ta stała się GP pewnej ZL X; naszkicować jej wykres.
b) Wyznaczyć dystrybuantę ZL X i naszkicować jej wykres.
c) Korzystając bezpośrednio z GP i wyznaczonej dystrybuanty obliczyć pr-stwa: $p_1 = P(X < 1)$, $p_2 = P(|X-1| < 0{,}5)$, $p_3 = P(X - 2X^2 > 0)$, $p_4 = P(\frac{1}{X} > 2)$; znalezione pr-stwa zilustrować na wykresie GP i wykresie dystrybuanty.

8. Dana jest funkcja f: $f(x) = \begin{cases} \frac{1}{2}\cos\frac{1}{2}x & \text{dla } 0 < x < \pi \\ 0 & \text{dla } x \leq 0 \vee x \geq \pi. \end{cases}$

a) Sprawdzić, że jest to GP pewnej ZL X, naszkicować jej wykres.
b) Wyznaczyć dystrybuantę ZL X i naszkicować jej wykres.
c) Obliczyć pr-stwa: $p_1 = P(X < \pi/2)$, $p_2 = P(X > \pi/3)$, $p_3 = P(\pi/3 < X < \pi/2)$ i zaznaczyć je na wykresie GP i dystrybuanty.

9. Dana jest GP f ZLC X. Wyznaczyć jej dystrybuantę F (wzór i wykres) oraz obliczyć pr-stwa p_k, jeśli:

a) $f(x) = \begin{cases} x^2/3 & \text{dla } -1 < x < 2, \\ 0 & \text{dla } x \leq -1 \vee x \geq 2; \end{cases}$ $p_1 = P(X < 1)$, $p_2 = P(0 < X < 1)$;

b) $f(x) = \begin{cases} 3(x-1)^2/2 & \text{dla } 0 < x < 2, \\ 0 & \text{dla } x \leq 0 \vee x \geq 2; \end{cases}$ $p_1 = P(|X-1| < 0{,}5)$, $p_2 = P(4X^2 - 8X + 3 > 0)$;

c) $f(x) = \begin{cases} x^3/4 & \text{dla } 0 < x < 2, \\ 0 & \text{dla } x \leq 0 \vee x \geq 2; \end{cases}$ $p_1 = P(|X-1| > 0{,}5)$, $p_2 = P(\frac{2X-1}{3-2X} > 0)$;

d) $f(x) = \begin{cases} 1/x^2 & \text{dla } x > 1, \\ 0 & \text{dla } x \leq 1; \end{cases}$ $p_1 = P(X < 3)$, $p_2 = P(X > 2)$;

e) $f(x) = \begin{cases} 2/x^3 & \text{dla } x > 1, \\ 0 & \text{dla } x \leq 1; \end{cases}$ $p_1 = P(4 - X^2 > 0)$, $p_2 = P(X^2 > 4X)$;

f) $f(x) = \begin{cases} 1/2\sqrt{x} & \text{dla } 1 < x < 4, \\ 0 & \text{dla } x \leq \vee x \geq 4; \end{cases}$ $p_1 = P(|X-3| < 1)$, $p_2 = P(X^3 - 4X < 0)$;

g) $f(x) = \begin{cases} 2e^{-2x} & \text{dla } x > 0 \\ 0 & \text{dla } x \leq 0; \end{cases}$ $p_1 = P(X > 1/X)$, $p_2 = P(X - 1 > 1/(X-1))$;

h) $f(x) = \begin{cases} \ln x & \text{dla } 1 < x < e, \\ 0 & \text{dla } x \leq 1 \vee x \geq e; \end{cases}$ $p_1 = P(X < \sqrt{e})$, $p_2 = P(X > e/X)$;

i) $f(x) = \begin{cases} 1-|x| & \text{dla } |x| < 1, \\ 0 & \text{dla } |x| \geq 1; \end{cases}$ $p_1 = P(|X| < 0{,}5)$, $p_2 = P(2X^2 - 5X + 2 < 0)$;

j) $f(x)=\begin{cases}1/2 \text{ dla } -1<x\le 0, & p_1=P(X>0),\\ x \text{ dla } 0<x<1, & p_2=P(4X^2-1>0),\\ 0 \text{ dla } |x|\ge 1; & p_3=P(|X|\ge 2X);\end{cases}$

k) $f(x)=\begin{cases}1/2 \text{ dla } 0<x\le 1, & p_1=P(0{,}5<X<1{,}5),\\ x-1 \text{ dla } 1<x<2, & p_2=P(\sin X>0{,}5),\\ 0 \text{ dla } x\le 0 \vee x\ge 2; & p_3=P(2^{2X}-3\sqrt{2}\,2^X+4>0);\end{cases}$

l) $f(x)=\begin{cases}x+1 \text{ dla } -1<x\le 0, & p_1=P(X^2<4),\\ 3x^2/2 \text{ dla } 0<x<1, & p_2=P(|X-1|<1),\\ 0 \text{ dla } |x|\ge 1; & p_3=P(|X|>0{,}5);\end{cases}$

ł) $f(x)=\begin{cases}\dfrac{1}{\pi\sqrt{4-x^2}} \text{ dla } |x|<2, & p_1=P(X>0),\\ 0 \text{ dla } |x|\ge 2; & p_2=P(|X|<1).\end{cases}$

10. Wyznaczyć GP f ZLC X mając daną jej dystrybuantę F:

a) $F(x)=\begin{cases}0 & \text{dla } x\le 1\\ 1-1/x & \text{dla } x>1,\end{cases}$

b) $F(x)=\begin{cases}-1/x & \text{dla } x\le -2\\ 1/2 & \text{dla } -2<x\le 2\\ 1-1/x & \text{dla } x>2,\end{cases}$

c) $F(x)=\begin{cases}0 & \text{dla } x\le -1\\ x/2 & \text{dla } -1<x\le 0\\ (x^2+1)/2 & \text{dla } 0<x\le 1\\ 1 & \text{dla } x>1,\end{cases}$

d) $F(x)=\frac{1}{2}+\frac{1}{\pi}\operatorname{arctg}3x,\ x\in R,$

e) $F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}dt,\ x\in R,$

f) $F(x)=\begin{cases}0 & \text{dla } x\le -a\\ \frac{1}{2}+\frac{1}{\pi}\arcsin\frac{x}{a} & \text{dla } -a<x\le a\\ 1 & \text{dla } x>a,\ a>0.\end{cases}$

11. Korzytając z twierdzenia 3.4 sprawdzić, czy następujące dystrybuanty F są dystrybuantami ZLC:

a)

x	$(-\infty,0>$	$(0,1>$	$(1,2>$	$(2,+\infty)$
F(x)	0	$x/2$	$x^2/2-x+1$	1

b) $F(x)=\begin{cases}0 & \text{dla } x\le 0\\ 1-e^{-x/2} & \text{dla } x>0,\end{cases}$

c) $F(x)=\begin{cases}0 & \text{dla } x\le 1\\ 0{,}3 & \text{dla } 1<x\le 5\\ 1 & \text{dla } x>5,\end{cases}$

d) $F(x)=\begin{cases}0 & \text{dla } x\le -2\\ \frac{1}{2}+\frac{1}{\pi}\arcsin\frac{x}{2} & \text{dla } -2<x\le 2\\ 1 & \text{dla } x>2,\end{cases}$

e) $F(x)=\begin{cases}0 & \text{dla } x\le 0\\ (x+2)/4 & \text{dla } 0<x\le 2\\ 1 & \text{dla } x>2.\end{cases}$

12. Dane są funkcje F:

(1) $F(x)=\begin{cases}0 & \text{dla } x\le 1\\ x\ln x-x+1 & \text{dla } 1<x\le e\\ 1 & \text{dla } x>e\end{cases}$

(2)

x	$(-\infty,1>$	$(1,e>$	(e,∞)
F(x)	0	0,5	1

(3) $F(x)=\begin{cases}0 & \text{dla } x\le 0\\ \cos x & \text{dla } 0<x\le \pi/2\\ 1 & \text{dla } x>\pi/2\end{cases}$

(4)

x	$(-\infty,0>$	$(0,1>$	$(1,\infty)$
F(x)	0	$\frac{1}{2}x+\frac{1}{4}$	1

a) Które z tych funkcji są dystrybuantami?

b) Które z nich są dystrybuantami ZLS? Wyznaczyć ich funkcje pr-stwa.

c) Które z nich są dystrybuantami ZLC? Podać postać GP.

13. Dobrać tak stałe a i b, aby funkcje postaci

$$F(x)=\begin{cases}0 & \text{dla } x\le 0\\ ax^2 & \text{dla } 0<x\le 2\\ 1 & \text{dla } x>2,\end{cases}\qquad G(x)=\begin{cases}0 & \text{dla } x\le 0\\ a+be^{-x} & \text{dla } x>0\end{cases}$$

były dystrybuantami: a) ZLS, b) ZLC.

14. ZLC X ma rozkład prostokątny skoncentrowany na przedziale (0,1), tj rozkład zadany GP postaci:

$$f(x)=\begin{cases}1 & \text{dla } 0<x<1\\ 0 & \text{dla } x\le 0 \vee x\ge 1\end{cases}$$

sprawdzić, że *różna od niej ZL* $Y\overset{\text{def}}{=}1-X$ *ma taką samą GP, a więc i taki sam rozkład pr-stwa.*

15. Funkcja $f_X(x)=\begin{cases}0 & \text{dla } x\le 0\\ e^{-x} & \text{dla } x>0\end{cases}$ jest GP ZL X. Wyznaczyć dystrybuantę i GP ZL Y, jeśli $Y=\sqrt{X}$.

Odpowiedzi

1. a) $c=0{,}3$, d) $p_1=0{,}1$, $p_2=0$, $p_3=0{,}5$, $p_4=1$, $p_5=0{,}8$, $p_6=0{,}4$, $p_7=0{,}5$, $p_8=0{,}6$, $p_9=0{,}9$.

2.

x_i	-3	0	4	9	15
p_i	0,2	0,2	0,3	0,1	0,2

3. a) $p_1=1-F(b)$, $p_2=F(b^+)$, $p_3=F(b^+)-F(a^+)$, $p_4=F(b^-)-F(a)$, $p_5=F(b)-F(a^+)$.

b) $p_1=1-F(b)$, $p_2=F(b)$, $p_3=p_4=p_5=F(b)-F(a)$.

c) $p_1=0{,}8$, $p_2=0{,}6$, $p_3=0{,}1$.

4.

y_j	0	1	2	3
p_j	1/8	3/8	3/8	1/8

X, Y są różnymi ZL, bo np. $X(\omega_1)=0\neq Y(\omega_1)=3$;

... jednakowe funkcje pr-stwa (a więc jednakowe rozkłady pr-stwa).

5.

x_i	9	11	13	15
p_i	1/35	12/35	18/35	4/35

6.

x_i	$\sqrt{3}/4$	$\sqrt{3}/2$	$3\sqrt{3}/4$
p_i	0,3	0,6	0,1

7. a) $a=1$.

b)

x	$(-\infty,0>$	$(0,1>$	$(1,2>$	$(2,\infty)$
F(x)	0	$\frac{1}{2}x^2$	$-\frac{1}{2}x^2+2x-1$	1

c) $p_1=0{,}5$, $p_2=0{,}75$, $p_3=1/8$, $p_4=p_3$.

8. b)

x	$(-\infty,0>$	$(0,\pi>$	(π,∞)
F(x)	0	$\sin\frac{1}{2}x$	1

c) $p_1=1/\sqrt{2}\approx 0{,}71$, $p_2=1/2$, $p_3=\frac{1}{\sqrt{2}}-\frac{1}{2}\approx 0{,}21$

9. a) $F(x)=\begin{cases}0 & \text{dla } x\le -1,\\ (x^3+1)/9 & \text{dla } -1<x\le 2,\\ 1 & \text{dla } x>2;\end{cases}$ $p_1=F(1)=2/9$, $p_2=F(1)-F(0)=1/9$.

b) $F(x)=\begin{cases}0 & \text{dla } x\le 0,\\ [(x-1)^3+1]/2 & \text{dla } 0<x\le 2,\\ 1 & \text{dla } x>2;\end{cases}$ $p_1=P(0{,}5<X<1{,}5)=1/8$, $p_2=1-p_1=7/8$.

c) $F(x)=\begin{cases}0 & \text{dla } x\le 0,\\ x^4/16 & \text{dla } 0<x\le 2,\\ 1 & \text{dla } x>2;\end{cases}$ $p_1=1-P(0{,}5<X<1{,}5)=5/16$, $p_2=1-p_1=11/16$.

d) $F(x)=\begin{cases}0 & \text{dla } x\le 1,\\ 1-1/x & \text{dla } x>1;\end{cases}$ $p_1=F(3)=2/3$, $p_2=1-F(2)=1/2$.

e) $F(x)=\begin{cases}0 & \text{dla } x\le 1,\\ 1-1/x^2 & \text{dla } x>1;\end{cases}$ $p_1=F(2)-F(-2)=3/4$, $p_2=P(X<0\vee X>4)=1/16$.

f) $F(x)=\begin{cases}0 & \text{dla } x\le 1,\\ \sqrt{x}-1 & \text{dla } 1<x\le 4,\\ 1 & \text{dla } x>4;\end{cases}$ $p_1=F(4)-F(2)=2-\sqrt{2}\approx 0{,}59$, $p_2=F(-2)+F(2)-F(0)=\sqrt{2}-1\approx 0{,}49$.

g) $F(x)=\begin{cases}0 & \text{dla } x\le 0,\\ 1-e^{-2x} & \text{dla } x>0;\end{cases}$ $p_1=P(X>1)=e^{-2}\approx 0{,}135$, $p_2=P(0<X<1)+P(X>2)=1-e^{-2}+e^{-4}\approx 0{,}883$.

h) $F(x)=\begin{cases}0 & \text{dla } x\le 1,\\ x\ln x - x + 1 & \text{dla } 1<x\le e,\\ 1 & \text{dla } x>e\end{cases}$ $p_1=F(e)=1-\sqrt{e}/2\approx 0{,}176,$

$p_2=P(X>\sqrt{e})=\sqrt{e}/2\approx 0{,}824.$

i)

x	$(-\infty,-1>$	$(-1,0>$	$(0,1>$	$(1,+\infty)$
F(x)	0	$(x+1)^2/2$	$(-x^2+2x+1)/2$	1

$p_1=3/4,$ $p_2=1/8.$

j)

x	$(-\infty,-1>$	$(-1,0>$	$(0,1>$	$(1,+\infty)$
F(x)	0	$(x+1)/2$	$(x^2+1)/2$	1

$p_1=p_3=1/2,$ $p_2=5/8.$

k)

x	$(-\infty,0>$	$(0,1>$	$(1,2>$	$(2,+\infty)$
F(x)	0	$x/2$	$x^2/2-x+1$	1

$p_1=3/8,\ p_2=P(\frac{\pi}{6}<X<\frac{5\pi}{6})=1-\pi/12\approx 0{,}738,\ p_3=1-p_1=5/8.$

l)

x	$(-\infty,-1>$	$(-1,0>$	$(0,1>$	$(1,+\infty)$
F(x)	0	$(x+1)/2$	$(x^3+1)/2$	1

$p_1=1,\ p_2=1/2,$ $p_3=9/16.$

ł) $F(x)=\begin{cases}0 & \text{dla } x\le -2,\\ \frac{1}{2}+\frac{1}{\pi}\arcsin\frac{x}{2} & \text{dla } -2<x\le 2,\\ 1 & \text{dla } x>2,\end{cases}$ $p_1=1/2,$ $p_2=1/3.$

10. Korzystamy z tw. 3.3, wzór (3.11).

a) $f(x)=\begin{cases}1/x^2 & \text{dla } x>1,\\ 0 & \text{dla } x\le 1;\end{cases}$

b) $f(x)=\begin{cases}1/x^2 & \text{dla } |x|>2,\\ 0 & \text{dla } |x|\le 2;\end{cases}$

c) $f(x)=\begin{cases}1/2 & \text{dla } -1<x\le 0,\\ x & \text{dla } 0<x<1,\\ 0 & \text{dla } |x|\ge 1;\end{cases}$

d) $f(x)=\frac{3}{\pi(9x^2+1)},\ x\in R;$

e) $f(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2},\ x\in R;$

f) $f(x)=\begin{cases}\frac{1}{\pi\sqrt{a^2-x^2}} & \text{dla } |x|<a,\\ 0 & \text{dla } |x|\ge a.\end{cases}$

11. a) tak, b) tak, c) nie, d) tak, e) nie.

12. a) (1), (2), (4),

b) (2).

x_i	1	e
p_i	0,5	0,5

c) (1), $f(x)=\begin{cases}\ln x & \text{dla } 1<x<e\\ 0 & \text{dla } x\le 1 \vee x\ge e;\end{cases}$ (4) jest dystrybuantą tzw. ZL skokowo-ciągłej.

13. a) a=0 dla F, a=1 i b=0 dla G; b) a=1/4 dla F, a=1 i b=-1 dla G.

15. $F_Y(y)=\begin{cases}0 & \text{dla } y\le 0\\ 1-e^{-y^2} & \text{dla } y>0,\end{cases}$ $f_Y(y)=\begin{cases}0 & \text{dla } y\le 0\\ 2ye^{-y^2} & \text{dla } y>0.\end{cases}$

4. CHARAKTERYSTYKI LICZBOWE ZMIENNYCH LOSOWYCH.

Zmiennym losowym przyporządkowuje się pewne liczby, które charakteryzują te zmienne pod różnymi względami, np. wielkości rozrzutu ich wartości. Liczby te nazywa się **charakterystykami liczbowymi** ZL.

Charakterystyka liczbowa ZL nazywa się **miarą położenia**, jeśli ma następującą własność: dodanie stałej do ZL powoduje dodanie tej stałej do tej charakterystyki. Charakterystykę liczbową nazywa się **miarą rozrzutu (rozproszenia)**, jeśli dodanie stałej do ZL nie zmienia wartości tej charakterystyki.

Ograniczymy się do omówienia dwu najważniejszych miar położenia - wartości oczekiwanej i kwantyli oraz dwu najważniejszych miar rozrzutu: wariancji i odchylenia standardowego. Inne charakterystyki, to m.in.: rozstęp, współczynnik asymetrii, odchylenie przeciętne, współczynnik zmienności, momenty.

WARTOŚĆ OCZEKIWANA. Poznawanie tej charakterystyki rozpoczniemy od prostego przykładu wprowadzającego nas w zagadnienie.

PRZYKŁAD 4.1. W sklepie znajduje się 10 sztuk towaru po cenie 2 zł za sztukę i 90 sztuk po cenie 8 zł za sztukę (od innego dostawcy). Interesuje nas przeciętna (oczekiwana) cena jednej sztuki, tj taka cena, przy której sklep uzyskałby za 100 sztuk towaru taki sam przychód jak przy faktycznych cenach. Zrozumiałe, że nie wynosi ona $(2+8)/2=5$ zł. Poprawną przeciętną cenę otrzymamy obliczając przychód za obydwa rodzaje towaru i dzieląc go przez 100:

$$\frac{2\cdot 10+8\cdot 90}{100}=2\cdot\frac{10}{100}+8\cdot\frac{90}{100}=2\cdot 0{,}1+8\cdot 0{,}9=7{,}4\text{ zł}.$$

Widzimy, że: *przeciętna cena uwzględnia nie tylko poszczególne ceny* $x_1=2$, $x_2=8$ *ale również częstości* $p_1=10/100$, $p_2=90/100$, *z jakimi te ceny występują.* ■

Wartością oczekiwaną (wartością przeciętną, wartością średnią) ZLS X o punktach skokowych x_i i skokach $p_i\equiv p(x_i)\equiv P(X=x_i)$ nazywamy liczbę EX określoną wzorem:

$$EX \overset{\text{def}}{=} \sum_{x_i\in W} x_i\, p(x_i) \tag{4.1}$$

pod warunkiem, że w przypadku przeliczalnej liczby punktów skokowych, szereg po prawej stronie jest bezwzględnie zbieżny.

Bezwzględna zbieżność zapewnia jednoznaczność określenia liczby EX. Jeśli bowiem szereg po prawej stronie wzoru jest tylko warunkowo zbieżny, to jego zbieżność, a w przypadku zbieżności - jego suma, zależą od porządku wyrazów (a więc od sposobu numeracji wartości ZL) i dlatego przyjmuje się, że w tym przypadku nie istnieje wartość oczekiwana ZL.

Wartość oczekiwana ma prostą interpretację fizyczną. Wzór (4.1), jeśli uwzględnić warunek unormowanie (3.2), można zapisać w postaci:

$$EX=\frac{\sum x_i\, p(x_i)}{\sum p(x_i)}$$

Wynika stąd, że: wartość oczekiwaną EX ZLS X o punktach skokowych x_i i skokach $p(x_i)$ można interpretować jako *współrzędną środka masy układu punktów materialnych* $M(x_i)$ *o masach* $m_i=p(x_i)$ (i całkowitej masie równej 1).

PRZYKŁAD 4.2. Wartość oczekiwana ZLS X o punktach skokowych $x_1=2$, $x_2=8$ i skokach odpowiednio $p_1=0{,}1$ i $p_2=0{,}9$ (por. przykład 4.1) jest "zwykłą" sumą:

$$EX=2\cdot 0{,}1+8\cdot 0{,}9=7{,}4.$$ ■

PRZYKŁAD 4.3. Wyznaczymy wartość oczekiwaną ZLS X przyjmującej wartości całkowite $x_k=k\in\{1,2,\dots\}$ z pr-stwami:

$$p_k\equiv P(X=k)=pq^{k-1},\quad k=1,2,\dots,$$

gdzie p jest daną liczbą z przedziału (0,1), zaś $q=1-p$. O takiej ZL mówimy, że ma **rozkład geometryczny z parametrem p**.

Zgodnie z definicją (4.1)

$$EX=\sum_{k=1}^{\infty} kpq^{k-1}=p\sum_{k=1}^{\infty} kq^{k-1}.$$

Mamy tu przeliczalną liczbę punktów skokowych, ale wszystkie są liczbami dodatnimi. Dlatego bezwzględna zbieżność szeregu po prawej stronie ostatniej równości jest równoważna zwykłej zbieżności. Szereg potęgowy można różniczkować "wyraz po wyrazie" (względem q), zatem:

$$EX=p\sum_{k=1}^{\infty} kq^{k-1}=p\sum_{k=1}^{\infty}(q^k)'=p\Big(\sum_{k=1}^{\infty} q^k\Big)'=p\Big(\frac{q}{1-q}\Big)'=p\frac{1}{(1-q)^2}=\frac{1}{p}.$$ ■

PRZYKŁAD 4.4 (ZL, dla której nie istnieje wartość oczekiwana). Chcemy wyznaczyć wartość oczekiwaną ZL X o funkcji pr-stwa postaci:

$$p_i\equiv P(X=x_i)=(1/2)^i,\quad \text{gdzie } x_i=(-2)^i/i,\ i=1,2,\dots.$$

Na mocy definicji

$$EX=\sum_{i=1}^{\infty} x_i p_i=\sum_{i=1}^{\infty}\frac{(-2)^i}{i}\frac{1}{2^i}=\sum_{i=1}^{\infty}(-1)^i\frac{1}{i}.$$

Szereg po prawej stronie jest zbieżny, ale nie bezwzględnie. Zatem, zgodnie z definicją, wartość oczekiwana EX nie istnieje. ■

Przy obliczaniu pr-stw zapewne Czytelnik zauważył, że sumowaniu dla ZLS odpowiada całkowanie w przypadku ZLC. Tak też jest przy określaniu wartości oczekiwanej.

Wartością oczekiwaną ZLC X o GP f nazywamy liczbę EX określoną wzorem:

$$EX \overset{\text{def}}{=} \int_{-\infty}^{\infty} x f(x)\,dx \tag{4.2}$$

pod warunkiem, że całka po prawej stronie jest bezwzględnie zbieżna (tzn. całka $\int_{-\infty}^{\infty}|x|f(x)\,dx$ jest zbieżna: $\int_{-\infty}^{\infty}|x|f(x)\,dx<\infty$).

Wzór (4.2), jeśli uwzględnić warunek unormowania (3.6) dla GP f, można zapisać w postaci:

$$EX=\int_{-\infty}^{\infty} x f(x)\,dx \Big/ \int_{-\infty}^{\infty} f(x)\,dx.$$

Zatem: jeśli GP f ZL X będziemy interpretować jako g ę s t o ś ć m a s y jednostkowej rozmieszczonej w sposób ciągły wzdłuż osi 0x, to *wartość oczekiwana* EX *jest współrzędną środka tej masy.*

P R Z Y K Ł A D 4.5.. Wyznaczymy wartość oczekiwaną EX ZLC X o **rozkładzie wykładniczym z parametrem** λ, tj. o rozkładzie zadanym GP postaci:

$$f(x)=\begin{cases}0 & \text{dla } x\le 0\\ \lambda e^{-\lambda x} & \text{dla } x>0,\end{cases}\qquad \lambda>0.$$

GP f jest tu skoncentrowana na dodatniej półosi, zatem zbieżność bezwzględna całki po prawej stronie (4.2) jest równoważna zwykłej zbieżności. Kolejno otrzymujemy:

$$EX=\int_{-\infty}^{\infty}xf(x)dx=\int_{0}^{\infty}x\lambda e^{-\lambda x}dx=\lambda\lim_{b\to\infty}\int_{0}^{b}xe^{-\lambda x}=\cdots=\frac{1}{\lambda}.$$ ■

Bardzo często zachodzi potrzeba wyznaczenia wartości oczekiwanej ZL $Y=g(X)$ będącej daną funkcją g ZL X, gdy dany jest rozkład pr-stwa ZL X. Wydawałoby się, że jedyna droga to: wyznaczenie najpierw funkcji pr-stwa ZL Y (gdy jest to ZLS) albo GP ZL Y (gdy jest to ZLC) i następnie zastosowanie wzorów (4.1) lub (4.2). Okazuje się, że nie jest to konieczne. Dowodzi się bowiem, że:

Wartość oczekiwana Eg(X) *wyraża się bezpośrednio przez funkcję pr-stwa* p(·) *lub* GP f ZL X *następującymi wzorami:*

$$Eg(X)=\sum_{x_i\in W}g(x_i)p(x_i), \tag{4.3}$$

$$Eg(X)=\int_{-\infty}^{\infty}g(x)f(x)dx \tag{4.4}$$

pod warunkiem, że szereg i całka po prawych stronach tych wzorów są bezwzględnie zbieżne.

W jednej z omawianych niżej własności wartości oczekiwanej mówi się o niezależnych zmiennych losowych. Sformułujemy teraz to pojęcie. Wiąże się ściśle ono z poznaną już niezależnością zdarzeń. Bardziej szczegółowo o niezależności ZL będziemy mówić w paragrafie 7.

Mówimy, że **ZL X i Y są niezależne**, gdy dla dowolnych $x,y\in R$ zdarzenia $\{\omega: X(\omega)<x\}$ i $\{\omega: Y(\omega)<y\}$ są niezależne, czyli

$$P(X<x, Y<y)=P(X<x)\cdot P(Y<y) \quad \text{dla dowolnych } x,y\in R. \tag{4.5}$$

Analogicznie określa się **niezależność** dowolnej skończonej liczby zmiennych losowych $X_1,X_2,\ldots,X_n$. Natomiast o zmiennych losowych nieskończonego ciągu $X_1,X_2,\ldots,X_n,\ldots$ powiemy, że są **niezależne**, gdy dla każdego ustalonego n niezależne są ZL $X_1,X_2,\ldots,X_n$.

Wartość oczekiwana ZL jest sumą lub całką. Można wobec tego oczekiwać, że ma ona własności podobne jak suma i całka. Zobaczymy, że istotnie tak jest.

TWIERDZENIE 4.1. Wartość oczekiwana ZL ma następujące własności (X,Y - ZL dowolnego typu, a, b, c - dowolne stałe): E1, E2, E3, E4, E5:

E1. Jeśli $P(X=c)=1$, to $EX=c$;
jeśli $P(X\ge 0)=1$, to $EX\ge 0$;
jeśli $P(a\le X\le b)=1$, to $a\le EX\le b$;
jeśli $P(X\le Y)=1$, to $EX\le EY$; w szczególności $EX\le E|X|$.

E2. Jeśli istnieje EX, to istnieją E(aX), E(X+b), E(aX+b)

oraz

$$E(aX)=aEX,\quad E(X+b)=EX+b,\quad E(aX+b)=aEX+b.$$

W szczególności: dla każdej ZL X, dla której istnieje EX, mamy:

$$E(X-EX)=0.$$

E3. Jeśli istnieją EX i EY, to istnieje $E(X\pm Y)$ oraz

$$E(X\pm Y)=EX\pm EY.$$

E4. Jeśli ZL X i Y są n i e z a l e ż n e oraz istnieją EX i EY, to istnieje E(XY) oraz

$$E(XY)=EX\cdot EY.$$

E5. Jeśli istnieje EX oraz krzywa gęstości (dla ZLC) lub histogram funkcji pr-stwa (dla ZLS) są symetryczne względem prostej $x=x_0$ (rozkład pr-stwa ZL X jest symetryczny względem punktu $x=x_0$), to $EX=x_0$.

D o w ó d . Ograniczymy się do dowodu pierwszej własności E1, trzeciej własności E2 oraz własności E5. Na mocy definicji EX, (wzór (4.1)), mamy:

$$EX=\sum_{x_i\in W}x_ip(x_i)=c\cdot P(X=c)=c\cdot 1=c.$$

Przechodzimy do dowodu trzeciej własności E2. Gdy X jest ZLS, to zgodnie z równością (4.3), w której przyjmujemy $g(X)=aX+b$ i warunkiem unormowania (3.2), mamy:

$$E(aX+b)=\sum_{x_i\in W}(ax_i+b)p(x_i)=a\sum_{x_i\in W}x_ip(x_i)+b\sum_{x_i\in W}p(x_i)=aEX+b.$$

Podobnie, gdy X jest ZLC o GP f, to na mocy równości (4.4) i warunku unormowania (3.6), mamy:

$$E(aX+b)=\int_{-\infty}^{\infty}(ax+b)f(x)dx=a\int_{-\infty}^{\infty}xf(x)dx+b\int_{-\infty}^{\infty}f(x)dx=aEX+b.$$

Ważna równość $E(X-EX)=0$ wynika z udowodnionej wyżej równości, jeśli przyjąć w tej równości $a=1$ i $b=-EX$.

Dowód własności E5 przeprowadzimy tylko dla ZLC. Z założenia EX istnieje, więc

$$EX=\int_{-\infty}^{\infty}xf(x)dx=\int_{-\infty}^{x_0}xf(x)dx+\int_{x_0}^{\infty}xf(x)dx.$$

Wykonując w całkach po prawej stronie odpowiednio podstawienia $x=x_0-t$, $x=x_0+t$, otrzymujemy:

$$EX=\int_{0}^{+\infty}(x_0-t)f(x_0-t)dt+\int_{0}^{+\infty}(x_0+t)f(x_0+t)dt.$$

Z założenia GP f jest symetryczna względem prostej $x=x_0$, więc $f(x_0-t)=f(x_0+t)$ dla każdego t oraz $\int_{x_0}^{+\infty}f(u)du=\frac{1}{2}$, zatem

$$EX=2x_0\int_{0}^{+\infty}f(x_0+t)dt=\left\{\begin{matrix}x_0+t=u\\dt=du\end{matrix}\right\}=2x_0\int_{x_0}^{+\infty}f(u)du=2x_0\cdot\frac{1}{2}=x_0.$$ □

KWANTYLE. Poznamy teraz charakterystykę liczbową ZL przydatną w następnym rozdziale. Ograniczymy się do ZLC. Niech zatem $p\in(0,1)$.

Kwantylem rzędu p ZLC X o dystrybuancie F i GP f nazywamy l i c z b ę x_p, spełniającą którykolwiek z następujących równoważnych warunków:

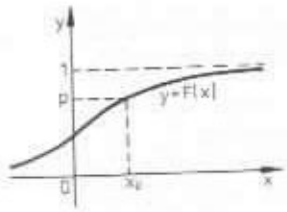
$$F(x_p)=p,\quad P(X<x_p)=p,\quad \int_{-\infty}^{x_p}f(x)dx=p. \tag{4.6}$$

Kwantyl rzędu $p=0{,}5$ nazywa się **medianą**.

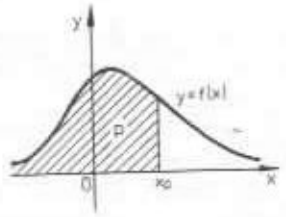
W interpretacji geometrycznej kwantyl x_p ZLC X o dystrybuancie F i GP f jest:

1) *odciętą* x_p punktu przecięcia prostej o równaniu $y=p$ z wykresem dystrybuanty $y=F(x)$, czyli pierwiastkiem równania $F(x)=p$, (por. rys 4.1),

2) *odciętą* x_p takiego punktu na osi 0x, że pole pod krzywą GP $y=f(x)$ nad przedziałem $(-\infty,x_p)$ jest równe p, (por. rys 4.2).



Rys 4.1.



Rys 4.2.

P R Z Y K Ł A D 4.6. Wyznaczymy kwantyl $x_{0,9}$ ZLC X z przykładu 4.5 dla $\lambda=2$.

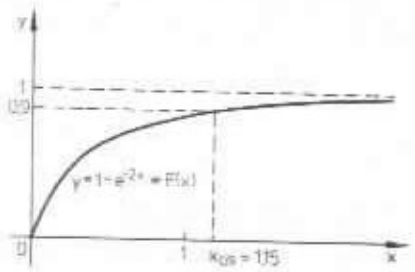
W znany sposób wyznaczamy najpierw dystrybuantę F ZL X:

$$F(x)=\begin{cases}0 & \text{dla } x\le 0\\ 1-e^{-\lambda x} & \text{dla } x>0.\end{cases}$$

Zrozumiałe, że $x_p\in(0,\infty)$. Zatem rozwiązując równanie

$F(x)=p$, czyli $1-e^{-\lambda x}=p$, otrzymujemy $x_p=-\frac{1}{\lambda}\ln(1-p)$.

W tym przykładzie udało się nam uzyskać jawny wzór na postać kwantyla. Przyjmując teraz $p=0{,}9$, $\lambda=2$ otrzymujemy: $x_{0,9}=-0{,}5\cdot\ln(1-0{,}9)=$ $=0{,}5\ln 10\approx 1{,}15$, (por. rys 4.3). ■



Rys 4.3.

WARIANCJA. Omówienie tej podstawowej miary rozproszenia rozpoczniemy od prostego przykładu wprowadzającego nas w zagadnienie.

PRZYKŁAD 4.7. Niech X, Y będą ZLS o funkcjach pr-stwa odpowiednio:

x_i	−2	2
$p(x_i)$	0,25	0,75

y_i	−20	15
$p(y_i)$	0,4	0,6

Mają one równe wartości oczekiwane: $EX=EY=1$. Różnią się jednak rozrzutem swych wartości względem punktu $x=1$ - druga ma bezsprzecznie większy rozrzut. Chcemy wprowadzić miarę tego rozrzutu. Miara taka powinna uwzględnić nie tylko "odchylenia" x_i-1 oraz y_i-1 ale również pr-stwa, z jakimi się one realizują. Jednak bezpośrednio odchylenia te nie mogą być wykorzystane, ponieważ dla każdej ZL Z zgodnie z ostatnią z własności E2, mamy $E(Z-EZ)=0$. Bierzemy wobec tego kwadraty tych odchyleń $(x_i-1)^2$, $(y_i-1)^2$ i obliczamy ich sumy mnożąc uprzednio przez pr-stwa $p(x_i)$ i $p(y_i)$ z jakimi się one realizują:

$$(-2-1)^2\cdot 0{,}25+(2-1)^2\cdot 0{,}75=3 \quad \text{oraz} \quad (-20-1)^2\cdot 0{,}4+(15-1)^2\cdot 0{,}6=294.$$

Tak określona miara rozrzutu dla ZL Y jest więc, zgodnie z pierwszym intuicyjnym osądem, znacznie większa od miary rozrzutu dla ZL X. ■

Realizując myśl zawartą w przykładzie przyjmujemy następujące określenie.

Wariancją ZL X dowolnego typu, dla której istnieje wartość oczekiwana EX, nazywamy l i c z b ę VarX (oznaczaną również symbolami: D^2X, σ^2, σ_x^2, μ_2) określoną wzorem

$$\text{VarX} \overset{\text{def}}{=} E(X-EX)^2 \tag{4.7}$$

Jest to więc *wartość oczekiwana kwadratu odchylenia* $X-EX$ ZL X *od jej wartości oczekiwanej* EX.

Uwzględniając w tej definicji równości (4.3) i (4.4) dla $g(X)=$ $=(X-EX)^2$, otrzymujemy:

$$\text{VarX}=\sum_{x_i\in W}(x_i-EX)^2p(x_i) \quad \text{dla ZLS X o skokach } p(x_i), \tag{4.8}$$

$$\text{VarX}=\int_{-\infty}^{\infty}(x-EX)^2f(x)dx \quad \text{dla ZLC X o GP f.} \tag{4.9}$$

Przeanalizujemy wzór (4.8) dla ZLS o skończonej liczbie punktów skokowych:

$$\text{VarX}=(x_1-EX)^2p(x_1)+(x_2-EX)^2p(x_2)+\cdots+(x_n-EX)^2p(x_n).$$

Wszystkie składniki po prawej stronie są nieujemne. Dlatego wariancja przyjmuje wartość bliską zera (jest mała) wtedy i tylko wtedy, gdy każdy składnik $(x_i-EX)^2p(x_i)$ jest bliski zera, tj gdy: 1) każdy punkt skokowy x_i mało różni się od EX lub 2) jeśli punkt skokowy x_i jest znacznie odległy od punktu EX, to pr-stwo $p(x_i)$ takiego zdarzenia jest bliskie zera. Zatem: *wariancja ZL jest mała jedynie wtedy, gdy zmienna ta przyjmuje wartości x_i ze zbioru blisko położonego wartości oczekiwanej* EX, *a jeśli zdarzają się wartości x_i ze zbioru znacznie odległego od punktu* EX, *to z pr-stwem $p(x_i)$ bliskim zera*. W tym sensie ***wariancja ZL jest miarą rozrzutu (rozproszenia) wartości ZL wokół jej wartości oczekiwanej: im mniejsza wariancja, tym rozkład ZL jest bardziej skupiony wokół swej wartości oczekiwanej; im większa wariancja, tym rozkład*** ZL ***jest bardziej rozproszony wokół swej wartości oczekiwanej.***

Sformułujemy teraz własności wariancji. Ułatwiają one obliczenie wariancji. Ponadto część z nich (własności: V1, V2, V4) pogłębiają nasze zrozumienie wariancji jako miary rozrzutu wartości ZL.

TWIERDZENIE 4.2. Wariancja ZL ma następujące własności (X,Y - ZL dowolnego typu, dla których istnieją wariancje, a, b, c - dowolne stałe) V1, V2, V3, V4, V5:

V1. Wariancja ZL jest liczbą nieujemną, $\mathrm{Var}X \geq 0$. Przy tym wariancja jest równa zero, $\mathrm{Var}X=0$, wtedy i tylko wtedy, gdy istnieje stała c taka, że $P(X=c)=1$, tj gdy ZL X ma **rozkład jednopunktowy**.

V2. $\mathrm{Var}(aX)=a^2\mathrm{Var}X$, $\mathrm{Var}(X+b)=\mathrm{Var}X$, $\mathrm{Var}(aX+b)=a^2\mathrm{Var}X$.

V3. $\mathrm{Var}X=E(X-c)^2-(c-EX)^2$, $\mathrm{Var}X=EX^2-(EX)^2$.

V4. $\mathrm{Var}X \leq E(X-c)^2$; równość zachodzi tylko dla $c=EX$.

V5. $\mathrm{Var}(X\pm Y)=\mathrm{Var}X+\mathrm{Var}Y$, gdy ZL X i Y są niezależne.

D o w ó d. Przeprowadzimy dowody tylko własności V2 i V3. Będziemy przy tym wykorzystywać poznane już własności wartości oczekiwanej.

Ad V2. $\mathrm{Var}(aX+b)=E[(aX+b)-E(aX+b)]^2=$ {własność E2} $=E[aX+b-aEX-b]^2=E[a(X-EX)]^2=E[a^2(X-EX)^2]=$

$=$ { własność E2 } $=a^2E[X-EX]^2=a^2\mathrm{Var}X$.

Pozostałe dwie równości otrzymujemy przyjmując odpowiednio $b=0$ oraz $a=1$.

Ad V3. Korzystając, jak dotychczas z odpowiednich własności wartości oczekiwanej i wykorzystując tożsamościowe przekształcenia, otrzymujemy:

$$\begin{aligned}\mathrm{Var}X&=E(X-EX)^2=E[(X-c)-(EX-c)]^2=\\&=E[(X-c)^2-2(EX-c)(X-c)+(EX-c)^2]=\\&=E(X-c)^2-2(EX-c)\cdot E(X-c)+E(EX-c)^2=\\&=E(X-c)^2-2(EX-c)\cdot(EX-c)+(EX-c)^2=\\&=E(X-c)^2-2(EX-c)^2+(EX-c)^2=E(X-c)^2-(c-EX)^2.\end{aligned}$$

Przyjmując teraz $c=0$ otrzymujemy drugą własność V3. □

U w a g a 1. Własności V3 dla ZLS przyjmują odpowiednio postać:

$$\mathrm{Var}X=\sum_{x_i\in W}(x_i-c)^2p_i-(c-EX)^2, \tag{4.10}$$

$$\mathrm{Var}X=\sum_{x_i\in W}x_i^2p_i-(EX)^2. \tag{4.11}$$

U w a g a 2. Przy obliczeniu wariancji ZLC najlepiej jest wykorzystać drugą własność V3 wariancji, która w tym przypadku przyjmuje postać:

$$\mathrm{Var}X=\int_{-\infty}^{\infty}x^2f(x)dx-(EX)^2. \tag{4.12}$$

P R Z Y K Ł A D 4.8. Obliczymy wariancję ZLS X, gdy:

a) jej funkcja pr-stwa jest postaci: (5;0,2), (10;0,4), (20;0,3), (25;0,1)

b) jej funkcja pr-stwa jest postaci (por. przykład 4.3):

$$p_k=P(X=k)=p(1-p)^{k-1},\quad k=1,2,\ldots .$$

a) Obliczamy przede wszystkim wartość oczekiwaną:

$$EX=\sum x_ip_i=5\cdot0{,}2+10\cdot0{,}4+20\cdot0{,}3+25\cdot0{,}1=13{,}5.$$

Aby mieć porównanie "uciążliwości rachunkowych" przy obliczaniu wariancji i aby ich uniknąć w przyszłości, wykorzystamy kolejno wszystkie trzy wzory (4.8), (4.10), (4.11):

$$\mathrm{Var}X=\{\text{ wzór (4.8) }\}=\sum_{i=1}^{4}(x_i-13{,}5)^2p_i=(5-13{,}5)^2\cdot0{,}2+(10-13{,}5)^2\cdot0{,}4+$$

$$+(20-13{,}5)^2\cdot0{,}3+(25-13{,}5)^2\cdot0{,}1=\cdots=42{,}25.$$

$$\mathrm{Var}X=\{\text{ wzór (4.10) dla c=14 }\}=\sum_{i=1}^{4}(x_i-14)^2p_i-(14-13{,}5)^2=$$

$$=9^2\cdot0{,}2+4^2\cdot0{,}4+6^2\cdot0{,}3+11^2\cdot0{,}1-(-0{,}5)^2=45{,}5-0{,}25=42{,}25.$$

$$\mathrm{Var}X=\{\text{ wzór (4.11) }\}=\sum_{i=1}^{4}x_i^2p_i-(13{,}5)^2=25\cdot0{,}2+100\cdot0{,}4+400\cdot0{,}3+$$

$$+625\cdot0{,}1-182{,}25=227{,}5-182{,}25=45{,}25.$$

Wydaje się, że w tym elementarnym przykładzie można zalecić stosowanie wzoru (4.10).

b) Wartość oczekiwaną $EX=1/p$ obliczyliśmy w przykładzie 4.3. Drugą własność V3 wariancji sprowadzamy do postaci:

$$\mathrm{Var}X=EX^2-(EX)^2=EX(X-1)+EX-(EX)^2=EX(X-1)+\frac{1}{p}-\frac{1}{p^2}.$$

Pzostaje wyznaczyć pierwszy składnik po prawej stronie. Postępujemy podobnie jak w przykładzie 4.3 z tym, że będziemy różniczkować dwukrotnie odpowiedni szereg potęgowy:

$$EX(X-1)=\sum_{k=1}^{\infty}k(k-1)pq^{k-1}=pq\sum_{k=2}^{\infty}k(k-1)q^{k-2}=pq\sum_{k=2}^{\infty}(q^k)''=$$

$$=pq(\sum_{k=2}^{\infty}q^k)''=pq(\frac{q^2}{1-q})''=pq\frac{2}{(1-q)^3}=\frac{2q}{p^2}.$$

Zatem

$$VarX=\frac{2q}{p^2}+\frac{1}{p}-\frac{1}{p^2}=\frac{q}{p^2}.$$ ■

PRZYKŁAD 4.9. Obliczymy wariancję ZL X o GP z przykładu 4.5. Obliczyliśmy tam wartość oczekiwaną $EX=1/\lambda$. Przy wyznaczaniu wariancji wykorzystuje się wzór (4.12), definicję całki niewłaściwej i regułę de l'Hospitala:

$$VarX=\int_{-\infty}^{\infty}x^2f(x)dx-(EX)^2=\lambda\int_0^{\infty}x^2e^{-\lambda x}dx-(EX)^2=\cdots=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}.$$

■

RÓŻNE UZUPEŁNIENIA. 1. ZL X w zastosowaniach jest modelem dla wielkości fizycznych. Gdy wielkość taka jest wyrażona w jednostkach "j" (np. w cm), to VarX wyraża się w jednostkach "j^2" (np. w cm^2). Jest to pewna wada wariancji. Wady tej nie ma **odchylenie standardowe** DX ZL X (oznaczane również σ,σ_x) określone wzorem

$$DX\overset{def}{=}\sqrt{VarX}. \tag{4.13}$$

Z odpowiednich własności wariancji wynika w szczególności, że:

(4.14) jeśli $P(X=c)=1$, to $DX=0$,

(4.15) $D(X+c)=DX$, $D(aX)=|a|DX$.

Z definicji, z własności (4.14) i pierwszej własności (4.15) wnioskujemy, że podobnie jak VarX, *odchylenie standardowe* DX *jest miarą rozrzutu* ZL X. Z samej definicji (4.13) wynika, że jest to ***miara rozrzutu* ZL X *wokół jej wartości oczekiwanej* EX.**

2. ZL Y dla której $EY=0$ i $DY=1$ ($VarY=1$) nazywa się **standaryzowaną zmienną losową**. Pozostawiamy Czytelnikowi sprawdzenie, że ZL Y określona wzorem

$$Y\overset{def}{=}\frac{X-EX}{DX} \tag{4.16}$$

jest standaryzowaną ZL. Czynności wykonywane na ZL X po prawej stronie wzoru (4.16) nazywa się **standaryzowaniem** ZL X.

3. Wartość oczekiwana i wariancja są szczególnymi przypadkami tzw. momentów ZL.

Momentem zwykłym rzędu r ZL X nazywa się liczbę α_r, określoną wzorem:

$$\alpha_r=EX^r,\quad r=0,1,2,\ldots. \tag{4.17}$$

Momentem centralnym rzędu r ZL X, dla której istnieje wartość oczekiwana EX, nazywa się liczbę μ_r określoną wzorem:

$$\mu_r=E(X-EX)^r. \tag{4.18}$$

Zgodnie z zapowiedzią mamy:

$$\alpha_1=EX,\quad \mu_2=VarX.$$

W związku z tym znaną nam dobrze równość $VarX=EX^2-(EX)^2$ można wyrazić w terminologii momentów:

$$\mu_2=\alpha_2-\alpha_1^2. \tag{4.19}$$

ZADANIA DO ROZWIĄZANIA.

1. Wyznaczyć wartość oczekiwaną i wariancję ZL rozważanych w zadaniach do paragrafu 3 o następujących numerach: a) 1, b) 2, c) 4, d) 5, e) 6.
2. Przeprowadzamy dwa identyczne doświadczenia D_1 i D_2 oraz w każdym z nich obserwujemy zajście tego samego zdarzenia A. Niech X oznacza różnicę między liczbą realizacji zdarzenia A i zdarzenia przeciwnego A' w tych dwóch doświadczeniach.
   a) Określić PZE Ω dla tego łącznego doświadczenia i opisaną już słownie ZL X opisać jako funkcję określoną na PZE Ω.
   b) Zakładając, że doświadczenia D_1 i D_2 są niezależne i pr-stwo zajścia zdarzenia A w każdym z nich jest jednakowe i wynosi $p=3/4$, wyznaczyć funkcję pr-stwa ZL X.
   c) Wyznaczyć wartość oczekiwaną i wariancję ZL X.

3. Wyznaczyć wartość oczekiwaną i wariancję ZL rozważanych w zadaniach do paragrafu 3 o następujących numerach: A) 7, B) 8, C) 9, D) 10.
4. Wyznaczyć wartość oczekiwaną ZL o **dwustronnym rozkładzie wykładniczym z parametrem** λ, $\lambda>0$, tj rozkładzie zadanym GP postaci: $f(x)=\frac{1}{2}\lambda e^{-\lambda|x|}$ dla $x\in R$.
5. Wyznaczyć wartość oczekiwaną i wariancję ZL: $Y=2X-1$, $U=2X^2-3$, $Z=-2X^2-3$, jeśli: $EX=1$, $VarX=2$, $EX^4=16$.
6. Wyznaczyć wartość oczekiwaną i wariancję ZL: $U=2X-3Y$, $Z=Y^2-X^2$, jeśli ZL X, Y są niezależne oraz $EX=2$, $VarX=3$, $EX^4=69$, $EY=-1$, $VarY=2$, $EY^4=14$.
7. Zbiory $\{EX, VarX, F(2), F(3)\}$, $\{1/3, 1/2, -3, 5\}$ są identyczne. Wyznaczyć wartości liczbowe elementów pierwszego zbioru.
8. Wyznaczyć kwantyle $x_{0,25}$, $x_{0,5}$, $x_{0,75}$ ZL z zadania 7 w paragrafie 3.
9. Wyznaczyć kwantyle $x_{0,25}, x_{0,5}, x_{0,75}$ ZL z zadania 4 przyjmując $\lambda=1$
10. Wykazać, że kwantyl jest miarą położenia (jeśli $Y=X+c$, to $y_p=x_p+c$).
11. Wykazać prawdziwość własności V5 wariancji sformułowanej w twierdzeniu 4.2.

Odpowiedzi.

1. a) $EX=6,3$, $VarX=8,61$; b) $EX=4,5$, $VarX=39,45$; c) $EY=3/2$, $VarY=3/4$; d) $EX=87/7$, $VarX=96/49$; e) $EX=9\sqrt{3}/20$, $VarX=27/400$.
2. a) $\Omega=\{(11),(10),(01),(0,0)\}$, gdzie $i_j=1$ lub 0 stosownie do tego, czy za j-tym ($j=1,2$) razem zrealizowało się zdarzenie A czy A'. $X(11)=2$, $X(10)=X(01)=0$, $X(00)=-2$; b) ZLS X ma skoki $p_1=1/16$, $p_2=6/16$, $p_3=9/16$ odpowiednio w punktach skokowych $x_1=-2$, $x_2=0$, $x_3=2$; c) $EX=1$, $VarX=3/2$.
3. A) $EX=1$, $VarX=1/6$; B) $EX=\pi-2$, $VarX=4(\pi-3)$,
   C) a) $EX=5/4$, $VarX=153/240\approx 0,64$;
   b) $EX=1$, $VarX=3/5$; c) $EX=8/5$, $VarX=8/75$;
   d) EX, VarX nie istnieją.; e) $EX=2$, VarX nie istnieje.;

   f) $EX=7/3$,, $VarX=34/45$; g) $EX=1/2$,, $VarX=1/4$;
   h) $EX=(e^2+1)/4$, $VarX=(-9e^4+32e^3-18e^2+7)/304$;
   i) $EX=0$, $VarX=1/6$; j) $EX=1/12$, $VarX=59/144\approx 0,41$;
   k) $EX=13/12$, $VarX=59/144\approx 0,41$;
   l) $EX=5/24$, $VarX=979/2880\approx 0,34$;
   l) $EX=0$, $VarX=2$.
D) a) EX, VarX nie istnieją; b) EX, VarX nie istnieją;
   c) $EX=1/12$, $VarX=59/144\approx 0,41$;
   d) EX, VarX nie istnieją;
   e) $EX=0$, $VarX=1$; f) $EX=0$, $VarX=a^2/2$.

4. $EX=0$; 5. $EY=1$, $VarY=8$; $EU=3$, $VarU=28$; $EZ=-9$, $VarZ=28$.
6. $EU=7$, $VarU=30$; $EZ=-4$, $VarZ=25$.
7. $F(2)=1/3$, $F(3)=1/2$, $EX=-3$, $VarX=5$.
8. $x_{0,25}=1/\sqrt{2}\approx 0,71$, $x_{0,5}=1$, $x_{0,75}=2-\sqrt{2}/2\approx 1,29$.
9. $x_{0,25}=-\ln 2\approx -0,69$, $x_{0,5}=0$, $x_{0,75}=\ln 2\approx 0,69$.

5. NIEKTÓRE SKOKOWE ROZKŁADY PRAWDOPODOBIEŃSTWA.

W tym i następnym paragrafie omówimy niektóre, najczęściej występujące rozkłady pr-stwa ZL. Rozpoczynamy od ZLS.

ZMIENNA O ROZKŁADZIE JEDNOPUNKTOWYM. Mówimy, że ZLS X ma **rozkład jednopunktowy**, jeżeli ma ona tylko jeden punkt skokowy x_1. Zatem jej funkcja pr-stwa $p(\cdot)$ i dystrybuanta F są postaci:

x_i	x_1
$p(x_i)$	1

x	$(-\infty,x_1>$	(x_1,∞)
F(x)	0	1

Wartość oczekiwana i wariancja tej zmiennej losowej wyrażają się wzorami (por. własności E1 *i* V1*):*

$$EX=x_1, \quad VarX=0.$$

ZL o rozkładzie jednopunktowym jest modelem dla wielkości, które po bliższej analizie okazują się wielkościami zdeterminowanymi; umożliwia traktowanie wielkości zdeterminowanych jako wielkości losowych.

ZMIENNA O ROZKŁADZIE ZERO-JEDYNKOWYM. Mówimy, że ZLS X ma **rozkład zero-jedynkowy z parametrem** p, $0<p<1$, co zapisujemy $X\sim(0-1)_p$, jeśli ma tylko dwa punkty skokowe $x_1=1$ i $x_2=0$ i skoki odpowiednio p oraz $q=1-p$. Zatem jej funkcja pr-stwa $p(\cdot)$ i dystrybuanta F są postaci:

x_i	0	1
$p(x_i)$	q	p

,

x	$(-\infty,0>$	$(0,1>$	$(1,\infty)$
F(x)	0	q	1

$p>0$
$q=1-p$

Wartość oczekiwana i wariancja ZL X *o rozkładzie* $(0-1)_p$ *wyrażają się wzorami:*

$$EX=p,\quad VarX=pq. \tag{5.1}$$

Równości te otrzymujemy na podstawie definicji:

$$EX=0\cdot q+1\cdot p=p,$$

$$VarX=(0-p)^2\cdot q+(1-p)^2\cdot p=p^2q+q^2p=pq(p+q)=pq.$$

ZL o rozkładzie $(0-1)_p$ jest modelem dla doświadczenia losowego, w którym wyniki klasyfikujemy jako "mające daną cechę W" z pr-stwem p i "nie mające tej cechy" z pr-stwem q. W szczególności jest użyteczna przy sprowadzaniu cech jakościowych do cech ilościowych.

PRZYKŁAD 5.1. Partia towaru składa się ze sztuk zgodnych z normą i pozostałych (wymagania mogą dotyczyć również cech jakościowych, np. koloru). Niech ZL X przyjmuje z pr-stwem p wartość 1, gdy sztuka towaru jest zgodna z normą, oraz z pr-stwem q wartość 0, gdy nie jest z nią w zgodzie. Tak określona ZL ma rozważany tu zero-jedynkowy rozkład pr-stwa z nieznanym parametrem p. ■

ZMIENNA O ROZKŁADZIE DWUMIANOWYM. Niech p oznacza daną liczbę z przedziału (0,1), n - daną liczbę naturalną, $n\in N$. Mówimy, że **ZLS S_n ma rozkład dwumianowy (binomialny, Bernoulliego) z parametrami (n,p)**, co zapisujemy $S_n\sim b(n,p)$, gdy jej punkty skokowe $x_k\equiv k$ tworzą zbiór $W=\{0,1,2,\ldots,n\}$ a skoki p_k, czyli pr-stwa $P(S_n=k)$, wyrażają się wzorem:

$$p_k\equiv P(S_n=k)=\binom{n}{k}p^kq^{n-k},\quad k=0,1,\ldots,n;\ 0<p<1,\ q=1-p. \tag{5.2}$$

Pr-stwa (5.2)są kolejnymi wyrazami rozwinięcia n-tej potęgi dwumianu $(q+p)^n$:

$$(q+p)^n=\binom{n}{0}p^0q^n+\binom{n}{1}p^1q^{n-1}+\cdots+\binom{n}{k}p^kq^{n-k}+\cdots+\binom{n}{n}p^nq^0,$$

stąd nazwa - rozkład dwumianowy. Ponadto z ostatniej równości, wobec $(q+p)^n=1^n=1$, widać, że pr-stwa p_k spełniają warunek unormowania:

$$p_0+p_1+\cdots+p_n=1.$$

Rozważana tu ZL S_n wiąże się ze znanym ze szkoły schematem Bernoulliego. Przypomnijmy: ciąg doświadczeń $D_1,D_2,\cdots,D_n$ nazywamy **schematem Bernoulliego (schematem n prób Bernoulliego)**, jeśli: 1) w każdym z tych doświadczeń może zajść zdarzenie A (zajście zdarzenia A nazywa się sukcesem, niezajście - porażką), 2) pr-stwo zajścia zdarzenia A (czyli pr-stwo sukcesu) w każdym pojedynczym doświadczeniu D_i jest jednakowe (powiedzmy p), 3) doświadczenia $D_1,D_2,\cdots,D_n$ są niezależne (w tym sensie, że wynik każdego doświadczenia D_i jest niezależny od wyników pozostałych doświadczeń).

A oto zapowiedziany związek: ZL S_n *o rozkładzie dwumianowym* b(n,p) *można interpretować jako możliwą liczbę sukcesów w schemacie n doświadczeń Bernoulliego z pr-stwem* p *sukcesu w pojedynczym doświadczeniu* D_i. Tym samym pr-stwa p_k dane wzorem (5.2) można interpretować jako *pr-stwa uzyskania k sukcesów w dowolnej kolejności w n takich doświadczeniach.*

Prawdziwa jest jeszcze jedna zależność. *ZL* S_n *o rozkładzie dwumianowym* b(n,p) *jest sumą n niezależnych ZL* $X_1,X_2,\cdots,X_n$ *o tym samym rozkładzie* $(0-1)_p$:

$$S_n=X_1+X_2+\cdots+X_n,\quad X_i\sim(0-1)_p \tag{5.3}$$

Istotnie. Niech bowiem dla $i=1,2,\ldots,n$, X_i oznacza ZL przyjmującą wartość 1 albo 0 stosownie do tego, czy w doświadczeniu D_i zrealizowało się zdarzenie A (sukces), czy też zdarzenie przeciwne A'

(porażka). Zatem wszystkie ZL X_i mają ten sam rozkład $(0-1)_p$. Wobec niezależności doświadczeń $D_1, D_2, \cdots, D_n$ - odpowiadające im ZL $X_1, X_2, \ldots X_n$ są również niezależne. Przy tych oznaczeniach dowodzona równość (5.3) wynika stąd, że zdania "ZL S_n przyjęła wartość k", "k spośród ZL $X_1, X_2, \ldots X_n$ przyjęło wartość 1" oznaczają to samo.

Związek (5.3) jest bardzo przydatny przy wyznaczaniu wartości oczekiwanej i wariancji ZL S_n. Korzystając bowiem z (5.3) i odpowiednich (jakich?) własności wartości oczekiwanej i wariancji oraz równości (5.1), otrzymujemy:

$$ES_n = E(X_1+X_2+\cdots+X_n) = EX_1+EX_2+\cdots+EX_n = p+p+\cdots+p = np,$$

$$VarS_n = Var(X_1+X_2+\cdots+X_n) = VarX_1+VarX_2+\cdots+VarX_n =$$

$$= pq+pq+\cdots+pq = npq.$$

Wykazaliśmy więc, że *wartość oczekiwana i wariancja* ZL S_n *o rozkładzie dwumianowym* b(n,p) *wyrażają się wzorami:*

(5.4) $$ES_n = np, \quad VarS_n = npq.$$

PRZYKŁAD 5.2. ZL X_1, X_2, X_3, X_4 są niezależne i mają ten sam rozkład $(0-1)_{0,6}$. Obliczymy pr-stwo tego, że ich suma przyjmie wartość: a) równą 2, b) co najmniej równą 2.

Suma $X_1+X_2+X_3+X_4$ jest ZL S_4 o rozkładzie b(4;0,6). Zatem:

a) $$P(X_1+X_2+X_3+X_4=2) = P(S_4=2) = \binom{4}{2}(0,6)^2(0,4)^{4-2} =$$

$$= 6\cdot 0,36\cdot 0,16 = 0,3456 \approx 0,35.$$

b) $$P(S_n \geq 2) = 1-P(S_n<2) = 1-[P(S_n=0)+P(X=1)] =$$

$$= 1-\binom{4}{0}(0,6)^0(0,4)^4-\binom{4}{1}(0,6)^1(0,4)^3 = 1-0,0256-0,1536 =$$

$$= 0,8208 \approx 0,82. \quad \blacksquare$$

ZMIENNA O ROZKŁADZIE POISSONA. Mówimy, że ZLS X ma **rozkład Poissona z parametrem** λ, $\lambda>0$, co zapisujemy: $X \sim Po(\lambda)$, jeśli jej funkcja pr-stwa jest postaci:

(5.5) $$p_k \equiv P(X=k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k=0,1,2,\ldots, \quad \lambda>0.$$

Zauważmy, że iloraz $\lambda^k/k!$ w (5.5) jest (k+1)-szym wyrazem rozwinięcia funkcji $f(\lambda)=e^{\lambda}$ w szereg Maclaurina:

$$e^{\lambda} = 1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\cdots+\frac{\lambda^k}{k!}+\cdots \equiv \sum_{k=0}^{\infty}\frac{\lambda^k}{k!}, \quad \lambda>0.$$

Mając to na uwadze łatwo widać, że pr-stwa (5.5) spełniają warunek unormowania:

$$\sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = e^0 = 1.$$

Wartość oczekiwana i wariancja ZL X o rozkładzie Poissona Po(λ) *wyrażają się wzorami:*

(5.6) $$EX=\lambda, \quad VarX=\lambda.$$

Ograniczymy się do uzasadnienia pierwszej równości:

$$EX = \sum_{x_k} x_k p_k = \sum_{k=0}^{\infty} k e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=1}^{\infty} k\frac{\lambda^k}{k!} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} =$$

$$= \lambda e^{-\lambda}\left(1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\cdots+\frac{\lambda^k}{k!}+\cdots\right) = \lambda e^{-\lambda}e^{\lambda} = \lambda. \quad \square$$

Obliczanie pr-stw p_k w rozkładzie dwumianowym b(n,p) jest uciążliwe. Obliczenie ich przybliżonych wartości za pomocą pr-stw p_k w rozkładzie Poissona umożliwia następujące

TWIERDZENIE 5.1 (o zbieżności rozkładu dwumianowego do rozkładu Poissona). Jeżeli ZL $S_1, S_2, \ldots, S_n, \ldots$ mają rozkłady dwumianowe $b(1,p_1), b(2,p_2), \ldots, b(n,p_n), \ldots$ oraz $np_n = const \equiv \lambda$, $\lambda>0$, to

(5.7) $$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k=0,1,2,\ldots.$$

Z tego twierdzenia wynika następujące **przybliżenie Poissona** rozkładu dwumianowego:

(5.8) $$\binom{n}{k}p^k(1-p)^{n-k} \approx e^{-\lambda}\frac{\lambda^k}{k!}, \quad \lambda=np, \quad k=0,1,2,\ldots,n.$$

Przybliżenie to jest wystarczająco dokładne, gdy: $p\leq 0,1$, $n\geq 50$, $np=\lambda\leq 10$.

PRZYKŁAD 5.3. W ramach wyrywkowej kontroli jakości produkcji z partii 2000 sztuk losujemy ze zwrotem 80 sztuk. Obliczymy przybliżoną wartość pr-stwa tego, że wśród tych wylosowanych sztuk znajdą się co najwyżej dwie sztuki wadliwe, jeśli w całej partii było 40 takich sztuk.

Mamy tu do czynienia ze schematem $n=80$ prób Bernoulliego. Jednak obliczanie pr-stw według wzoru Bernoulliego dla $n=80$ byłoby bardzo kłopotliwe. Dlatego posłużymy się przybliżeniem Poissona (5.8). Mamy tu:

$$p=\frac{40}{2000}=0{,}02<0{,}1,\quad n=80\geq 50,\quad \lambda=np=80\cdot 0{,}02=1{,}6\leq 10,$$

$$P(S_n\leq 2)=P(S_n=0)+P(S_n=1)+P(S_n=2)\approx e^{-1{,}6}\frac{(1{,}6)^0}{0!}+$$

$$+e^{-1{,}6}\frac{(1{,}6)^1}{1!}+e^{-1{,}6}\frac{(1{,}6)^2}{2!}=3{,}88\cdot e^{-1{,}6}\approx 0{,}78.$$

Zatem interesujące nas pr-stwo wynosi 0,78. ■

ZMIENNA O SKOKOWYM ROZKŁADZIE RÓWNOMIERNYM. Mówimy, że ZLS X **ma skokowy rozkład równomierny**, jeśli ma ona skończoną liczbę punktów skokowych $x_1<x_2<\ldots<x_n$ i wszystkie skoki równe. Jej funkcja pr-stwa i dystrybuanta są zatem postaci:

x_i	x_1	x_2	.....	x_n
p_i	$1/n$	$1/n$		$1/n$

x	$(-\infty,x_1>$	$(x_1,x_2>$	$(x_2,x_3>$	...	$(x_{n-1},x_n>$	$(x_n,\infty>$
F(x)	0	$1/n$	$2/n$		$(n-1)/n$	1

Wartość oczekiwana i wariancja ZL X o skokowym rozkładzie równomiernym wyrażają się wzorami:

$$(5.9)\qquad EX=\frac{1}{n}\sum_{i=1}^{n}x_i,\qquad VarX=\frac{1}{n}\sum_{i=1}^{n}(x_i-EX)^2.$$

ZADANIA DO ROZWIĄZANIA.

1. Wyznaczyć maksymalną wartość wariancji ZL o rozkładzie: a) zero-jedynkowym z parametrem p, b) dwumianowym b(n,p).

2. W stałych warunkach pr-stwo odebrania sygnału radiowego wynosi 0,8. Wyznaczyć pr-stwo, że pięciokrotnie nadany sygnał zostanie odebrany co najmniej dwa razy.
3. ZL X i Y mają rozkłady dwumianowe odpowiednio b(4;0,5) i b(8;0,5). Która z liczb $P(X=3)$, $P(Y=5)$ jest większa?
4. ZL $X_1,\ldots,X_5$ są niezależne i mają ten sam rozkład zero-jedynkowy z parametrem $p=1/5$. a) Znaleźć rozkłady ZL $S_k=X_1+\ldots+X_k$ $k=2,3,4,5$. b) Obliczyć pr-stwo $P(X_1+X_3+X_5=1)$.
5. Na ten sam obiekt wystrzelono serię 5 pocisków. Pr-stwo trafienia celu jednym pociskiem wynosi 0,8. Niech X oznacza liczbę trafień celu. Obliczyć $P(X\geq 3)$.
6. Z partii towaru o wadliwości 1% wylosowano ze zwrotem 200 sztuk towaru. Obliczyć pr-stwo, że wśród wylosowanych sztuk znajdą się: a) dwie wadliwe sztuki towaru, b) co najwyżej dwie wadliwe sztuki towaru.

Odpowiedzi.

1. a) $1/4$, b) $n/4$. 2. X - liczba odebranych sygnałów, $X\sim b(5;0{,}8)$, $P(X\geq 2)=1-P(X=0)-P(X=1)\approx 0{,}41$. 3. $P(X=3)=0{,}25>P(Y=5)\approx 0{,}22$.
4. a) $S_k\sim b(k;1/5)$, $k=2,3,4,5$, b) $S_3\sim b(3;1/5)$, $P(S_3=1)=\binom{3}{1}(0{,}2)^1(0{,}8)^2=0{,}384$.
5. $X\sim b(5;0{,}8)$, $P(X\geq 3)=0{,}2048+0{,}4096+0{,}32768\approx 0{,}94$.
6. $X\sim Po(2)$, a) $e^{-2}2^2/2!\approx 0{,}27$, b) $P(X\leq 2)=P(X=0)+P(X=1)+P(X=2)\approx 0{,}67$

6. NIEKTÓRE CIĄGŁE ROZKŁADY PRAWDOPODOBIEŃSTWA.

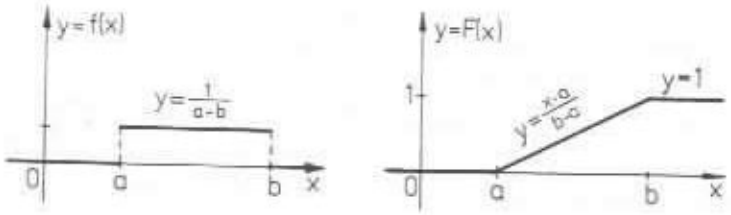
ZMIENNA O ROZKŁADZIE PROSTOKĄTNYM. Mówimy, że ZLC X ma **rozkład prostokątny (równomierny, jednostajny)** na przedziale (a,b), co zapisujemy $X\sim R(a,b)$, jeśli jej GP f wyraża się wzorem (rys 6.1):

$$(6.1)\qquad f(x)=\begin{cases}\dfrac{1}{b-a} & \text{dla } a<x<b\\ 0 & \text{dla } x\leq a \vee x\geq b.\end{cases}$$

Stąd łatwo wyznaczamy (wyznaczyć) dystrybuantę F(x), wartość oczekiwaną EX i wariancję VarX ZL X o rozkładzie równomiernym R(a,b):

$$F(x)=\begin{cases}0 & \text{dla } x\le a\\ \dfrac{x-a}{b-a} & \text{dla } a<x\le b\\ 1 & \text{dla } x>0,\end{cases}$$

(6.2) $$EX=\frac{1}{2}(a+b),\quad VarX=\frac{1}{12}(b-a)^2.$$



Rys 6.1.

Jeżeli uwzględnimy, że rozkład prostokątny R(a,b) można interpretować w fizyce jako ciągły rozkład masy jednostkowej na przedziale (a,b) ze stałą gęstością liniową masy $\rho(x)=1/(b-a)$, to wzory (6.2) stają się oczywiste. Pierwszy z nich, to współrzędna środka masy jednorodnego pręta pokrywającego przedział (a,b), drugi - to moment bezwładności względem osi prostopadłej do pręta i przechodzącej przez jego środek.

Zmienna losowa o rozkładzie prostokątnym R(a,b) jest:

1) modelem dla zagadnień, w których można założyć, że pr-stwo otrzymania wyniku obserwacji w danym przedziale $(x_1,x_2)\subset(a,b)$ jest proporcjonalne do długości x_2-x_1 tego przedziału i nie zależy od jego położenia w przedziale (a,b);

2) ciągłym odpowiednikiem ZL o skokowym rozkładzie równomiernym (gdzie jak wiemy, wszystkie punkty skokowe x_i są jednakowo prawdopodobne).

ZMIENNA O ROZKŁADZIE WYKŁADNICZYM. Mówimy, że ZLC ma rozkład **wykładniczy z parametrem** λ, $\lambda>0$, co zapisujemy $X\sim Ex(\lambda)$, gdy jej GP jest postaci (rys. 3.9 dla $\lambda=3$):

(6.3) $$f(x)=\begin{cases}0 & \text{dla } x\le 0\\ \lambda e^{-\lambda x} & \text{dla } x>0.\end{cases}$$

Stąd otrzymujemy postać jej dystrybuanty (rys. 3.8 dla $\lambda=3$):

$$F(x)=\begin{cases}0 & \text{dla } x\le 0\\ 1-e^{-\lambda x} & \text{dla } x>0\end{cases}$$

oraz wartość oczekiwaną i wariancję (por. przykłady 4.5 i 4.9):

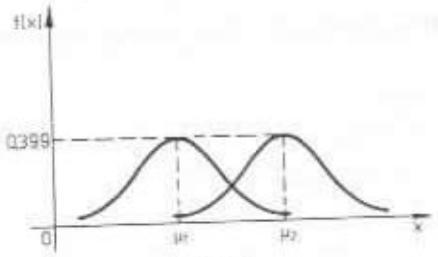
(6.4) $$EX=1/\lambda,\quad VarX=1/\lambda^2.$$

Zmienna losowa o rozkładzie wykładniczym $Ex(\lambda)$ jest w teorii niezawodności modelem dla czasu bezawaryjnej pracy między dwoma kolejnymi uszkodzeniami.

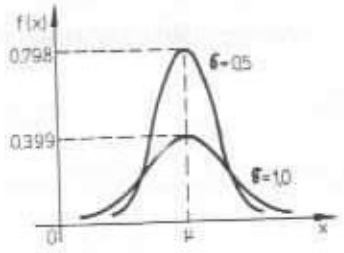
ZMIENNA O ROZKŁADZIE NORMALNYM. Mówimy, że ZLC X ma **rozkład normalny (rozkład Gaussa)** z parametrami (μ,σ), μ - dowolne, $\sigma>0$, co zapisujemy $X\sim N(\mu,\sigma)$, gdy jej GP f jest postaci

(6.5) $$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}},\quad -\infty<x<\infty.$$

Wykres tej GP nazywa się **krzywą normalną** lub **krzywą Gaussa**. Wpływ parametrów μ i σ na kształt i położenie krzywej normalnej ilustrują rysunki 6.2 i 6.3.



Rys 6.2.



Rys 6.3.

Krzywa normalna ma następujące własności: 1) jest symetryczna względem prostej $x=\mu$, 2) dla $x=\mu$ ma maksimum równe $1/(\sigma\sqrt{2\pi})\approx 0{,}4/\sigma$, 3) oś 0x jest jej asymptotą poziomą w $\pm\infty$, 4) $x_1=\mu-\sigma$, $x_2=\mu+\sigma$ są odciętymi jej punktów przegięcia, 5) jest wypukła ku górze w przedziale $(\mu-\sigma,\mu+\sigma)$, dla pozostałych x jest wypukła ku dołowi.

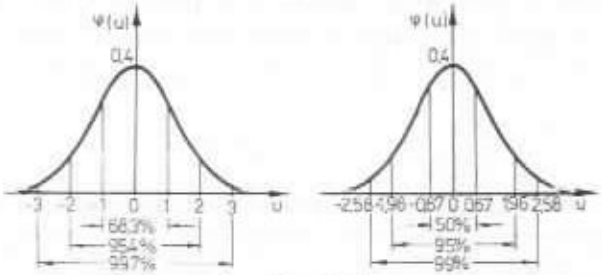
Parametry μ i σ mają następującą interpretację probabilistyczną: *μ jest wartością oczekiwaną, σ jest odchyleniem standardowym* ZL X *o rozkładzie* $N(\mu,\sigma)$:

(6.6) $$\mu=EX,\quad \sigma=DX,\ \text{czyli } \sigma^2=VarX.$$

ZL U o rozkładzie normalnym z wartością oczekiwaną $\mu=0$ i odchyleniem standardowym $\sigma=1$ nazywa się zmienną losową o **standaryzowanym rozkładzie normalnym** N(0,1). Jej GP, którą będziemy oznaczać literą φ, jest więc postaci:

$$\varphi(u)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}, \quad -\infty<u<\infty. \tag{6.7}$$

Krzywa gęstości ZL o rozkładzie N(0,1) i związane z nią wartości niektórych czarakterystycznych pr-stw są pokazane na rysunku 6.4.



Rys 6.4.

Istnieją tablice wartości GP φ i odpowiadającej jej dystrybuanty Φ

$$\Phi(u)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u}e^{-\frac{1}{2}t^2}dt \tag{6.8}$$

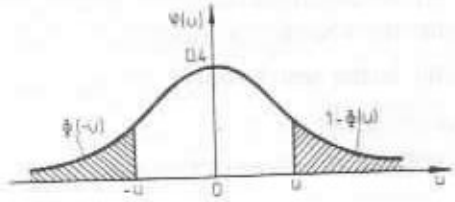
oraz tablice kwantyli, tj liczb u_p spełniających warunek

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u_p}e^{-\frac{1}{2}t^2}dt=p, \quad \text{czyli} \quad \Phi(u_p)=p.$$

Tablice wartości dystrybuanty Φ są sporządzone dla $u\geq 0$. Obliczanie wartości $\Phi(u)$ dla $u<0$ umożliwia następująca własność dystrybuanty Φ (rys. 6.5)::

$$\Phi(-u)=1-\Phi(u) \quad \text{dla każdego } u\in R. \tag{6.9}$$

Prawdziwość tej równości wynika z następujących przesłanek: 1) krzywa $y=\varphi(u)$ jest symetryczna względem osi rzędnych, 2) zakreskowane pola na rysunku 6.5 są równe, 3) pole po lewej stronie jest równe $\Phi(-u)$, po prawej stronie to $1-\Phi(u)$.



Rys 6.5.

Tablice kwantyli u_p ZL U o rozkładzie N(0,1) są sporządzane dla $0,5\leq p<1$. Znalezienie kwantyli u_p ZL U o rozkładzie N(0,1) dla $0<p<0,5$ umożliwia następująca ich własność:

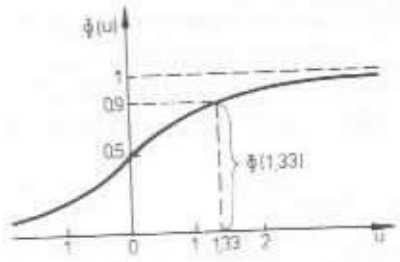
$$u_p=-u_{1-p}. \tag{6.10}$$

Aby uzasadnić tę równość zastosujemy najpierw własność (6.9) dla $u=-u_p$:

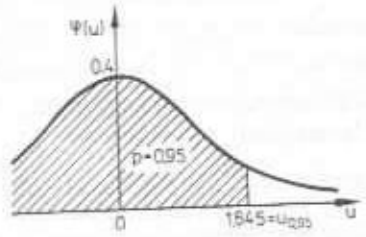
$$\Phi(-u_p)=1-\Phi(u_p)=1-p, \quad \text{czyli} \quad \Phi(-u_p)=1-p.$$

Z ostatniej równości ponownie na podstawie definicji kwantyla danego rzędu, wnioskujemy że liczba $-u_p$ jest kwantylem rzędu 1-p ZL U:

$$-u_p=u_{1-p}, \quad \text{czyli} \quad u_p=-u_{1-p}.$$



Rys 6.6.



Rys 6.7.

PRZYKŁAD 6.1. ZL U ma rozkład N(0,1). Wyznaczymy: a) $P(U<1,33)$, b) liczbę u spełniającą warunek $P(U<u)=0,95$.

a) szukane pr-stwo jest wartością dystrybuanty Φ rozkładu N(0,1) dla $u=1,33$, (por.rys 6.6):

$$P(U<1,33)=\Phi(1,33)=0,9082\approx 0,91,$$

wartość tę odczytujemy z tablicy Ia, znajdującej się na końcu książki.

b) szukaną liczbą jest kwantyl: $u=u_{0,95}=1,645$, (por. rys 6.7), odczytujemy go z tablicy Ib. ■

TWIERDZENIE 6.1. Jeżeli ZL X ma rozkład normalny $N(\mu,\sigma)$, to ZL U postaci

$$U \overset{\text{def}}{=} \frac{X-\mu}{\sigma} \tag{6.11}$$

ma standaryzowany rozkład normalny N(0;1).

D o w ó d . Należy sprawdzić, że GP φ ZL U jest postaci (6.7). Niech F i f oznaczają dystrybuantę i GP ZL X o rozkładzie $N(\mu,\sigma)$. Wyrazimy najpierw dystrybuantę ZL U za pomocą dystrybuanty F ZL X:

$$\Phi(u)=P(U<u)=P(\frac{X-\mu}{\sigma}<u)=P(X<\mu+\sigma u)=F(\mu+\sigma u).$$

Różniczkując teraz dystrybuantę Φ względem zmiennej u, otrzymamy GP φ ZL U:

$$\varphi(u)=\Phi'(u)=[F(\mu+\sigma u)]'=F'(\mu+\sigma u)\cdot\sigma=\sigma f(\mu+\sigma u)=$$

$$=\frac{\sigma}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(\mu+\sigma u-u)^2}{2\sigma^2}}=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}. \qquad \square$$

Przekształcenie (6.11) dokonane na ZL X nazywa się, jak wiemy, standaryzowaniem ZL X. Ostatnie twierdzenie umożliwia więc, drogą standaryzacji ZL X o rozkładzie $N(\mu,\sigma)$, wyznaczanie różnych pr-stw dotyczących ZL X za pomocą dystrybuanty Φ ZL U o rozkładzie N(0,1). Na przykład:

$$P(a<X<b)=\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma}); \tag{6.12}$$

w szczególności dla $a=-\infty$, $b=+\infty$ mamy

$$P(X<b)=\Phi(\frac{b-\mu}{\sigma}), \quad P(X>a)=1-\Phi(\frac{a-\mu}{\sigma}). \tag{6.13}$$

Równość (6.12) otrzymujemy w rezultacie następującego ciągu równości:

$$P(a<X<b)=P(\frac{a-\mu}{\sigma}<\frac{X-\mu}{\sigma}<\frac{b-\mu}{\sigma})=$$

$$=P(\frac{a-\mu}{\sigma}<U<\frac{b-\mu}{\sigma})=\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma}).$$

P R Z Y K Ł A D 6.2. ZL X ma rozkład N(5,2). Wyznaczymy: a) $P(X<5)$, b) $P(4<X<5)$, c) $P(|X|<4)$, d) $P(|X-4|\geq 2)$, e) $P(3X-X^2>0)$, f) $P(\frac{5}{X}<1)$.

a) Krzywa gęstości ZL X jest symetryczna względem prostej $x=5$, zatem, bez korzystania z tablic, mamy: $P(X<5)=0,5$.

b) Kolejno na mocy (6.12) i (6.9) otrzymujemy:

$$P(4<X<5)=\Phi(\frac{5-5}{2})-\Phi(\frac{4-5}{2})=\Phi(0)-\Phi(-0,5)=$$

$$=0,5-[1-\Phi(0,5)]=0,5-1+0,6915=0,1915\approx 0,19;$$

c) $P(|X|<4)=P(-4<X<4)=\Phi(\frac{4-5}{2})-\Phi(\frac{-4-5}{2})=$

$$=\Phi(-0,5)-\Phi(-4,5)=1-\Phi(0,5)-0=1-0,6915=0,3085\approx 0,31;$$

d) $P(|X-4|\geq 2)=1-P(|X-4|<2)=1-P(-2<X-4<2)=$

$$=1-P(2<X<6)=1-[\Phi(0,5)-\Phi(-1,5)]=1-[\Phi(0,5)-1+\Phi(1,5)]=$$

$$=2-\Phi(0,5)-\Phi(1,5)=2-0,6915-0,9332=0,3755\approx 0,38;$$

e) $P(3X-X^2>0)=P(0<X<3)=\Phi(-1)-\Phi(-2,5)=$

$$=1-\Phi(1)-1+\Phi(2,5)=0,9938-0,8413=0,1525\approx 0,15;$$

f) $P(\frac{5}{X}<1)=P(X<0 \text{ albo } X>5)=P(X<0)+P(X>5)=$

$$=\Phi(-2,5)+0,5=1-\Phi(2,5)+0,5=1,5-0,9938=0,5062\approx 0,51.$$

■

ZADANIA DO ROZWIĄZANIA.

1. Uzasadnić prawdziwość wzorów (6.2) na wartość oczekiwaną i wariancję ZL o rozkładzie prostokątnym. Wyznaczyć dystrybuantę F tej ZL.
2. Autobusy z przystanku miejskiej linii odjeżdżają co 8 minut. Chwila przyjścia pasażera na przystanek może być traktowana jako wartość ZL. Sensownie jest założyć, że ta ZL ma rozkład prostokątny. Przy tym założeniu obliczyć pr-stwo, że pasażer będzie czekał na autobus co najwyżej 3 minuty.
3. ZL T o GP postaci $f(t)=\begin{cases}0 & \text{dla } t\leq 0\\ \frac{1}{5}e^{-t/5} & \text{dla } t>0\end{cases}$ jest dobrym modelem dla czasu bezawaryjnej pracy pewnego urządzenia. Obliczyć:

a) przeciętny czas bezawaryjnej pracy urządzenia, b) pr-stwo, że urządzenie będzie pracować bezawaryjnie co najmniej przez 10 jednostek czasu.

4. ZL X ma rozkład wykładniczy o wartości oczekiwanej $EX=1/2$. Wyznaczyć: a) kwantyle: $x_{0,25}$, $x_{0,5}$, $x_{0,75}$,

b) pr-stwa: $p_1=P(X>1)$, $p_2=P(\frac{X}{X+1}<\frac{1}{2})$, $p_3=P(\frac{X}{X+1}<1)$.

5. ZL U ma rozkład N(0;1) o dystrybuancie Φ.

a) Sprawdzić, że $P(|U|<a)=2\Phi(a)-1$,

b) Wyznaczyć liczbę b spełniającą warunek $P(U<b)=0,99$,

c) Wyznaczyć liczbę b spełniającą warunek $P(U>a)=0,95$,

d) Wyznaczyć liczby a i b spełniające warunek $P(a<U<b)=0,95$,

e) Wyznaczyć najkrótszy przedział (a,b) spełniający warunek $P(a<U<b)=0,95$,

f) Wyznaczyć najdłuższy przedział (a,b) spełniający warunek $P(a<U<b)=0,95$.

6. ZL X ma rozkład N(3,2). Wyznaczyć pr-stwa:
$p_1=P(X<3)$, $p_2=P(1<X<3)$, $p_3=P(|X-3|<1)$,
$p_4=P(|X|>1)$, $p_5=P(X^2-X>0)$.
Podać interpretację geometryczną znalezionych pr-stw na wykresie GP ZL X.

7. ZL X ma rozkład $N(\mu,\sigma)$. Sprawdzić, że

$$P(|X-\mu|<k\sigma)=P(\mu-k\sigma<X<\mu+k\sigma)=\begin{cases}0,6827 & \text{gdy } k=1\\ 0,9545 & \text{gdy } k=2\\ 0,9973 & \text{gdy } k=3\end{cases}.$$

Ostatnia równość jest podstawą sformułowania tzw. **reguły trzysigmowej**: *jeżeli* ZL X *ma rozkład* $N(\mu,\sigma)$, *to zdarzenie "zmienna ta przyjmie wartość z przedziału* $(\mu-3\sigma,\mu+3\sigma)$*" jest praktycznie pewne.*

8. Dokonując pomiaru popełniamy z nadmiarem błąd systematyczny 1 cm i błąd losowy X. Zakładając, że X ma rozkład N(0;0,3) obliczyć: a) wartość przeciętną całkowitego błędu pomiaru, b) pr-stwo, że co do modułu całkowity błąd nie przekracza 1,5 cm.

Odpowiedzi.

2. $P(X\le 3)=0,375$. 3. a) $ET=1/\lambda=5$, b) $P(T\ge 10)=e^{-2}\approx 0,14$.

4. a) $x_p=\frac{-1}{\lambda}\ln(1-p)$, $x_{0,25}\approx 0,14$, $x_{0,5}\approx 0,35$, $x_{0,75}\approx 0,69$; b) $p_1=e^{-2}\approx 0,14$, $p_2=1-e^{-2}\approx 0,86$, $p_3=1$.

5. b) $b=u_{0,99}=2,33$; c) $a=u_{0,05}=-u_{0,95}=-1,64$; d) jest nieskończenie wiele przedziałów spełniających dany warunek, np. przedziały z punktów e) i f);
e) (-1,96;1,96), (por. rys 6.4); f) $(-1,64;\infty)$ lub $(-\infty;1,64)$.

6. $p_1=0,5$; $p_2=0,34$; $p_3=0,38$; $p_4=0,86$; $p_5=0,91$.

8. a) $E(X+1)=1$; b) $P(|X+1|\le 1,5)=\Phi(1,67)-\Phi(-8,33)\approx 0,95$.

7. WEKTORY LOSOWE.

Rozważane do tej pory jednowymiarowe ZL są modelami dla doświadczeń, których poszczególny wynik może być opisany jedną liczbą rzeczywistą x. Modelami dla doświadczeń, w których opis pojedynczego wyniku wymaga ciągu $(x_1,x_2,\ldots,x_n)$ liczb rzeczywistych są n-wymiarowe wektory losowe.

WEKTORY LOSOWE I ICH ROZKŁADY PRAWDOPODOBIEŃSTWA. Ciąg $(X_1,X_2,\ldots,X_n)$ jednowymiarowych ZL $X_1,X_2,\ldots,X_n$ określonych na wspólnej PP $(\Omega,\mathcal{A},P)$ nazywa się **n-wymiarowym wektorem losowym** (WL) lub **n-wymiarową ZL**.

Niech dla ustalonego $\omega\in\Omega$ $X_1(\omega)=x_1, X_2(\omega)=x_2,\ldots, X_n(\omega)=x_n$. Zatem: n-wymiarowy WL $(X_1,X_2,\ldots,X_n)$ każdemu zdarzeniu elementarnemu $\omega\in\Omega$ przyporządkowuje punkt $(x_1,x_2,\ldots,x_n)\in R^n$.

Ograniczamy się do omówienia dwuwymiarowych WL (X,Y). Wektor taki każdemu zdarzeniu elementarnemu $\omega\in\Omega$ przyporządkowuje parę (x,y) liczb rzeczywistych, czyli punkt na płaszczyźnie R^2.

Gdy $A\subset R^2$, to zapis $(X,Y)\in A$ odczytujemy: "WL (X,Y) przyjmie wartość ze zbioru A". Z relacjami postaci $(X,Y)\in A$ chcemy skojarzyć pr-stwo P ich realizacji (bardziej poprawnie byłoby oznaczyć to pr-stwo P innym symbolem, np. $P_{(X,Y)}$). Przyjmujemy:

(7.1) $$P((X,Y)\in A) \overset{\text{def}}{=} P\{\omega: (X(\omega),Y(\omega))\in A\}$$

czyli: *pr-stwo* $P((X,Y)\in A)$ *przyjęcia przez* WL (X,Y) *wartości ze zbioru płaskiego* A, *określamy jako pr-stwo zdarzenia składającego się z tych zdarzeń elementarnych* ω, *dla których* WL (X,Y) *przyjmuje wartość* $(X(\omega),Y(\omega))$ *ze zbioru* A.

Gdy dla każdego zbioru płaskiego $A\subset R^2$ dane jest pr-stwo $P((X,Y)\in A)$, to mówimy, że dany jest **rozkład pr-stwa WL** (X,Y).

Wartość (x_i,y_j) WL (X,Y) nazywa się jego **punktem skokowym** WL (X,Y), jeśli pr-stwo $P(X=x_i,Y=y_j)$, oznaczane przez $p(x_i,y_j)$ lub p_{ij}, jest dodatnie:

$$P(X=x_i,Y=y_j)\equiv p(x_i,y_j)\equiv p_{ij}>0;$$

przy tym liczba p_{ij} nazywa się **skokiem WL** (X,Y) w punkcie skokowym (x_i,y_j). *Zbiór* W *punktów skokowych* WL *może być skończony albo przeliczalny.*

Podobnie jak w przypadku jednowymiarowym wyróżnia się dwa podstawowe typy WL: WL skokowe (dyskretne) i WL ciągłe.

WEKTORY LOSOWE SKOKOWE. WL (X,Y) o zbiorze W jego punktów skokowych (x_i,y_j) nazywa się **wektorem losowym skokowym (WLS)**, gdy suma jego skoków $p_{ij}=P(X=x_i,Y=y_j)$ spełnia warunek unormowania, tj:

(7.2) $$\sum_{(i,j)} p_{ij}=1.$$

Jeśli (X,Y) jest WLS, to funkcję $p(\cdot,\cdot)$ określoną na zbiorze W jego punktów skokowych (x_i,y_j) równością

(7.3) $$p(x_i,y_j) \overset{\text{def}}{=} P(X=x_i,Y=y_j)\equiv p_{ij},\quad (x_i,y_j)\in W$$

lub tabelą (7.5) (bez ostatniego wiersza i ostatniej kolumny) nazywa się **funkcją (rozkładu) pr-stwa WLS** (X,Y).

Gdy $A\subset R^2$, to na mocy definicji przyjmujemy:

(7.4) $$P((X,Y)\in A) \overset{\text{def}}{=} \sum_{(x_i,y_j)\in A} p(x_i,y_j)\equiv \sum_{(x_i,y_j)\in A} p_{ij}.$$

Widać stąd, że *funkcja pr-stwa* $p(\cdot,\cdot)$ *WLS* (X,Y) *w pełni określa jego rozkład pr-stwa* $P((X,Y)\in A)$.

(7.5)

y_j \ x_i	x_1	x_2	...	x_i	...	$p_{\cdot j}$
y_1	p_{11}	p_{21}	...	p_{i1}	...	$p_{\cdot 1}$
y_2	p_{12}	p_{22}	...	p_{i2}	...	$p_{\cdot 2}$
...	...	...	...	...	...	...
y_j	p_{1j}	p_{2j}	...	p_{ij}	...	$p_{\cdot j}$
...	...	...	...	...	...	...
$p_{i\cdot}$	$p_{1\cdot}$	$p_{2\cdot}$	...	$p_{i\cdot}$	...	1

PRZYKŁAD 7.1. Interesuje nas zgodność z normą cech A i B u każdego z pięciu elementów $\omega_1,\omega_2,\ldots,\omega_5$. Dostawcy wiadomo, że: 1) element ω_1 nie odpowiada wymaganiom normy tylko ze względu na cechę A, 2) element ω_2 nie odpowiada wymaganiom normy tylko ze względu na cechę B, 3) element ω_3 nie odpowiada wymaganiom normy ze względu na obie cechy A i B, 4) pozostałe elementy są zgodne z normą. Doświadczenie polega na wylosowaniu jednego elementu i zanotowaniu obu jego cech A i B. Mamy tu: $\Omega=\{\omega_1,\omega_2,\ldots,\omega_5\}$, $\mathcal{A}$ - rodzina wszystkich 32 podzbiorów zbioru Ω, pr-stwo P określamy zgodnie z klasycznym modelem: $P(\{\omega_i\})=1/5$ dla $i=1,2,\ldots,5$. Niech ZL X przyjmuje wartość 1, gdy element spełnia wymagania normy ze względu na cechę A i wartość 0 w sytuacji przeciwnej; analogicznie określamy ZL Y w przypadku cechy B.

WL (X,Y) określony jest więc następująco:

$(X(\omega_1),Y(\omega_1))=(0,1)$, $(X(\omega_2),Y(\omega_2))=(1,0)$, $(X(\omega_3),Y(\omega_3))=(0,0)$, $(X(\omega_4),Y(\omega_4))=(X(\omega_5),Y(\omega_5))=(1,1)$.

Zbiorem punktów skokowych jest zbiór czteroelementowy: $W=\{(0,0),(0,1),(1,0),(1,1)\}$. Skoki p_{ij} w tych punktach obliczamy zgodnie z (7.1). Na przykład:

$p_{12}=P(X=0,Y=1)=P\{\omega_i: (X(\omega_i),Y(\omega_i))=(0,1)\}=P(\{\omega_1\})=1/5=0,2$,

$p_{22}=P(X=1,Y=1)=P\{\omega_i: (X(\omega_i),Y(\omega_i))=(1,1)\}=P(\{\omega_4,\omega_5\})=2/5=0,4$

Podobnie dla $A=\{(x,y)\in R^2: x\le 1, y\ge 1\}$, na mocy definicji (7.5), obliczamy:

$$P(X\le 1,Y\ge 1)=P((X,Y)\in A)=\sum_{(x_i,y_j)\in A} p_{ij}=p_{12}+p_{22}=0,2+0,4=0,6.$$

Funkcję pr-stwa $p(\cdot,\cdot)$ rozważanego WL (X,Y) można opisać tabelą (bez ostatniego wiersza i bez ostatniej kolumny) postaci (7.5):

y_j \ x_i	0	1	p_j
0	0,2	0,2	0,4
1	0,2	0,4	0,6
p_i	0,4	0,6	1

■

WEKTORY LOSOWE CIĄGŁE. Mówimy, że WL (X,Y) jest (dwuwymiarowym) **wektorem losowym ciągłym (WLC)**, gdy istnieje nieujemna i całkowalna na całej płaszczyźnie 0xy funkcja f taka, że dla każdego prostokąta $D=\{(x,y): x_1\le x\le x_2,\ y_1\le y\le y_2\}$

$$(7.6)\qquad P(x_1\le X\le x_2,\ y_1\le Y\le y_2)=\iint_D f(x,y)dxdy=\int_{x_1}^{x_2}\left[\int_{y_1}^{y_2} f(x,y)dy\right]dx;$$

przy tym funkcję f nazywa się **gęstością pr-stwa (GP) WL** (X,Y), a jej wykres w układzie 0xyz - **powierzchnią gęstości**. Z (7.6) wynika, że

$$(7.7)\qquad \iint_{R^2} f(x,y)dxdy\equiv\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x,y)dy\right]dx=1 \quad \text{(warunek unormowania)}$$

gdyż, zgodnie z (7.6), lewa strona (7.7) jest pr-stwem zdarzenia pewnego. Z (7.7) wynika, że *objętość bryły ograniczonej powierzchnią gęstości i płaszczyzną 0xy jest równa jedności.*

Dla $x_1=x_2$ lewa strona (7.6) jest pr-stwem przyjęci przez WLC (X,Y) wartości ze zbioru punktów odcinka $\{(x,y): x=x_1,\ y_1\le y\le y_2\}$, prawa zaś strona jest równa zero. Zatem: *pr-stwo przyjęcia przez* WLC *wartości ze zbioru punktów dowolnego odcinka* (ogólniej: *ze zbioru punktów będących wykresem funkcji* $y=g(x)$ *lub* $x=h(y)$) *jest zawsze równe zero.*

GP f WLC (X,Y) można interpretować jako **gęstość masy** jednostkowej na płaszczyźnie 0xy.

PRZYKŁAD 7.2. GP f WLC (X,Y) jest postaci:

$$f(x,y)=\begin{cases}2xy & \text{dla } (x,y)\in T\\ 0 & \text{dla pozost.}(x,y),\end{cases}$$

gdzie T jest trójkątem określonym nierównościami $0<y\le x<\sqrt{2}$. Obliczymy pr-stwo:

$$P(X\ge 1,Y\ge 1)=\int_1^{\sqrt{2}}\left[\int_1^x 2xydy\right]dx=\int_1^{\sqrt{2}}(x^3-x)dx=\frac{1}{4}.$$ ■

ROZKŁADY BRZEGOWE. Rozkład pr-stwa WL (X,Y) wyznacza rozkłady pr-stwa ZL X i ZL Y. Wynika to z następujących dwóch twierdzeń.

TWIERDZENIE 7.1. Jeżeli (X,Y) jest WLS o punktach skokowych (x_i,y_j) i skokach p_{ij} oraz

$$(7.8)\qquad p_i \overset{\text{def}}{=} \sum_j p_{ij}, \qquad p_j \overset{\text{def}}{=} \sum_i p_{ij},$$

to współrzędne X i Y są jednowymiarowymi ZLS o funkcjach pr-stwa odpowiednio postaci:

$$(7.9)\qquad P(X=x_i)=p_i \quad \textit{oraz} \quad P(Y=y_j)=p_j.$$

D o w ó d. Ograniczymy się do ZL X. Prawdziwość pierwszej równości (7.9) wynika z następującego ciągu równości:

$$P(X=x_i)=P(X=x_i\wedge Y \text{ przyjmie dowolną wartość})=$$
$$=P[X=x_i\wedge(Y=y_1\vee Y=y_2\vee\cdots\vee Y=y_j\vee\cdots)]=$$
$$=P[(X=x_i,Y=y_1)\vee(X=x_i,Y=y_2)\vee\cdots\vee(X=x_i,Y=y_j)\vee\cdots]=$$
$$=P[(X=x_i,Y=y_1)+P(X=x_i,Y=y_2)+\cdots+P(X=x_i,Y=y_j)+\cdots=$$
$$=\sum_j p_{ij}=p_i$$ □

TWIERDZENIE 7.2. Jeżeli (X,Y) jest WLC o GP f, to jego współrzędne X i Y są jednowymiarowymi ZLC o GP odpowiednio postaci:

$$(7.10)\qquad f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy, \qquad f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx.$$

Powszechnie przyjęła się następująca terminologia.

1) Funkcje pr-stwa $p_i=P(X=x_i)$ dla ZL X i $p_j=P(Y=y_j)$ dla ZL Y nazywa się **brzegowymi funkcjami pr-stwa WLS** (X,Y), (por. tabela (7.5)).

2) GP f_X i f_Y nazywa się **brzegowymi GP WLC** (X,Y).

3) Dystrybuanty F_X i F_Y odpowiednio ZL X i ZL Y nazywa się **brzegowymi dystrybuantami WL (X,Y)**.

4) Rozkłady pr-stwa ZL X i ZL Y wyznaczane przez brzegowe funkcje pr-stwa, brzegowe GP lub brzegowe dystrybuanty nazywa się **brzegowymi rozkładami pr-stwa WL** (X,Y).

PRZYKŁAD 7.3. Brzegowe funkcje pr-stwa WLS (X,Y) z przykładu 7.1 zawiera pierwszy i ostatni wiersz dla ZL X oraz pierwsza i ostatnia kolumna dla ZL Y zamieszczonej tam tabeli określającej funkcję pr-stwa WLS (X,Y):

x_i	0	1
p_i	0,4	0,6

y_j	0	1
p_j	0,4	0,6

ZL X i Y mają więc rozkłady zero-jedynkowe z parametrem $p=0,6$.

PRZYKŁAD 7.4. Rozważamy WLC (X,Y) o GP f z przykładu 7.2. Wyznaczymy brzegowe GP f_X i f_Y. Korzystamy z (7.10). Dla $y\le 0 \vee y\ge\sqrt{2}$ $f_Y(y)=0$. Dla $0<y<\sqrt{2}$

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx=\int_{-\infty}^{y} 0\,dx+\int_{y}^{\sqrt{2}} 2xy\,dx+\int_{\sqrt{2}}^{\infty} 0\,dx=yx^2\Big|_y^{\sqrt{2}}=y(2-y^2).$$

Zatem

$$f_Y(y)=\begin{cases} y(2-y^2) & \text{dla } 0<y<\sqrt{2}\\ 0 & \text{dla } y\le 0 \vee y\ge\sqrt{2}.\end{cases}$$

Analogicznie znajdujemy: $f_X(x)=\begin{cases} x^3 & \text{dla } 0<x<\sqrt{2}\\ 0 & \text{dla } x\le 0 \vee x\ge\sqrt{2}.\end{cases}$ ■

ROZKŁADY WARUNKOWE. Rozważmy WL (X,Y). Interesuje nas wpływ informacji o jednej z tych ZL na pr-stwa związane z drugą ZL. Niech $B\subset R$ oraz $P(Y\in B)>0$. Wówczas dla zbiorów $A\subset R$, zgodnie z (2.1), określamy pr-stwo warunkowe

$$(7.11)\qquad P(X\in A|Y\in B)=\frac{P(X\in A, Y\in B)}{P(Y\in B)}.$$

Wzór ten określa pewien rozkład pr-stwa na prostej generowany rozkładem P wektora losowego (X,Y). Nazywamy go **warunkowym rozkładem pr-stwa ZL X przy warunku, że ZL Y przyjęła wartość** y **ze zbioru** B.

W szczególności, gdy (X,Y) jest WLS o punktach skokowych (x_i,y_j) i skokach p_{ij}, $A=\{x_i\}$, $B=\{y_j\}$, to otrzymujemy:

$$(7.12)\qquad P(X=x_i|Y=y_j)=\frac{P(X=x_i, Y=y_j)}{P(Y=y_j)}=\frac{p_{ij}}{p_j}.$$

Wzór ten określa **warunkową funkcję pr-stwa** $P(X=x_i|Y=y_j)$ **ZLS X przy warunku, że ZLS Y przyjęła wartość** y_j. Analogicznie określa się **warunkową funkcję pr-stwa** $P(Y=y_j|X=x_i)$ **ZL Y przy warunku, że ZL X przyjęła wartość** x_i.

Niech teraz (X,Y) będzie WLC o GP f. Interesuje nas pr-stwo (7.11), gdy $A=(-\infty,x)$ i $B=\{y\}$, czyli $P(X<x|Y=y)$. Jednak $P(Y=y)=0$, gdyż Y jest ZLC. Definicja (7.11) jest więc w tym przypadku nieprzydatna. Przyjmujemy (jeśli rozważana granica istnieje):

$$(7.13)\qquad P(X<x|Y=y)\overset{\text{def}}{=}\lim_{h\to 0+}P(X<x|y\le Y<y+h)\equiv$$
$$\equiv\lim_{h\to 0+}\frac{P(X<x, y\le Y<y+h)}{P(y\le Y<y+h)}$$

przy założeniach, że $P(y\le Y<y+h)>0$ oraz $y\in\{y: f_Y(y)>0\}$. Pr-stwo $P(X<x|Y=y)$, określone równością (7.13) nazywa się **dystrybuantą warunkową ZL X przy warunku** $Y=y$ i oznacza $F_X(x|y)$ lub bardziej poprawnie $F_X(\cdot|y)$. GP odpowiadającą dystrybuancie $F_X(\cdot|y)$ nazywa się **warunkową GP ZL X przy warunku** $Y=y$ i oznacza $f_X(\cdot|y)$.

TWIERDZENIE 7.3. Jeżeli (X,Y) jest WLC o GP f, ciągłej względem każdego z argumentów x, y oraz $y\in\{y: f_Y(y)>0\}$, $x\in\{x: f_X(x)>0\}$, to warunkowe GP $f_X(\cdot|y)$ i $f_Y(\cdot|x)$ są postaci:

(7.14) $f_X(x|y)=f(x,y)/f_Y(y)$, $f_Y(y|x)=f(x,y)/f_X(x)$.

PRZYKŁAD 7.5. Dla WLS (X,Y) z przykładu 7.1 istnieją dwie warunkowe funkcje pr-stwa dla ZL X: $P(X=x_i|Y=0)$, $P(X=x_i|Y=1)$ i dwie dla ZL Y: $P(Y=y_j|X=0)$, $P(Y=y_j|X=1)$. Np. dla ZL Y otrzymujemy:

$$P(Y=0|X=1)=\frac{P(X=1,Y=0)}{P(X=1)}=\frac{p_{21}}{p_{2\cdot}}=\frac{0,2}{0,6}=\frac{1}{3},$$

$$P(Y=1|X=1)=\frac{P(X=1,Y=1)}{P(X=1)}=\frac{p_{22}}{p_{2\cdot}}=\frac{0,4}{0,6}=\frac{2}{3},$$

y_j	0	1
$P(Y=y_j\|X=1)$	1/3	2/3

Jest to rozkład zero-jedynkowy z parametrem $p=2/3$. Informacja o tym, że ZL X przyjęła wartość 1 zmieniła parametr rozkładu ZL Y z $p=3/5$ na $p=2/3$.

Analogicznie wyznacza się pozostałe warunkowe funkcje pr-stwa:

x_i	0	1
$P(X=x_i\|Y=0)$	1/2	1/2

x_i	0	1
$P(X=x_i\|Y=1)$	1/3	2/3

y_j	0	1
$P(Y=y_j\|X=0)$	1/2	1/2

■

PRZYKŁAD 7.6. Dla WLC (X,Y) z przykładu 7.4 wyznaczymy warunkowe GP $f_X(\cdot|y)$ i $f_Y(\cdot|x)$.

Warunkowa GP ZL X $f_X(\cdot|y)$ istnieje dla tych y, dla których brzegowa GP $f_Y(y)>0$, tj. dla $y\in(0,\sqrt{2})$. Niech zatem $y\in(0,\sqrt{2})$. Wówczas:

1) dla $x\le y \vee x\ge\sqrt{2}$ $f_X(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{0}{y(2-y^2)}=0$;

2) dla $y<x<\sqrt{2}$ $f_X(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{2xy}{y(2-y^2)}=\frac{2x}{2-y^2}$.

Zatem warunkowa GP $f_X(\cdot|y)$ istnieje dla $y\in(0,\sqrt{2})$ i wyraża się wzorami:

$$f_X(x|y)=\begin{cases}\frac{2x}{2-y^2} & \text{dla } y<x<\sqrt{2}\\ 0 & \text{dla } x\le y \vee x\ge\sqrt{2}.\end{cases}$$

Analogicznie: warunkowa GP $f_Y(\cdot|x)$ istnieje dla $x\in(0,\sqrt{2})$ i wyraża się wzorami:

$$f_Y(y|x)=\begin{cases}\frac{2y}{x^2} & \text{dla } 0<y<x\\ 0 & \text{dla } y\le 0 \vee y\ge x.\end{cases}$$

■

LINIE REGRESJI PIERWSZEGO RODZAJU. Rozkładom warunkowym można przyporządkować **warunkowe wartości oczekiwane** $E(Y|x_i)$ i $E(X|y_j)$ w przypadku dyskretnym oraz $E(Y|x)$ i $E(X|y)$ w przypadku ciągłym, określone równościami:

(7.15) $E(Y|x_i)\overset{\text{def}}{=}\sum_{y_j} y_j P(Y=y_j|X=x_i)$, $E(X|y_j)\overset{\text{def}}{=}\sum_{x_i} x_i P(X=x_i|Y=y_j)$,

(7.16) $E(Y|x)\overset{\text{def}}{=}\int_{-\infty}^{\infty} y f_Y(y|x)dy$, $E(X|y)\overset{\text{def}}{=}\int_{-\infty}^{\infty} x f_X(x|y)dx$.

PRZYKŁAD 7.7. a) Dla WLS (X,Y) z przykładu 7.5 mamy:

$E(Y|0)=0\cdot\frac{1}{2}+1\cdot\frac{1}{2}=\frac{1}{2}$, $E(Y|1)=0\cdot\frac{1}{3}+1\cdot\frac{2}{3}=\frac{2}{3}$, $E(X|0)=\frac{1}{2}$, $E(X|1)=\frac{2}{3}$.

b) Dla WLC (X,Y) z przykładu 7.6 mamy:

$$E(Y|x)=\int_0^x y\frac{2y}{x^2}dy=\frac{2}{3}\frac{y^3}{x^2}\Big|_0^x=\frac{1}{3}x \quad \text{dla } x\in(0,\sqrt{2}),$$

$$E(X|y)=\int_y^{\sqrt{2}} x\frac{2x}{2-y^2}dx=\frac{2}{3}\frac{2\sqrt{2}-y^3}{2-y^2}=\frac{2}{3}\frac{y^2+y\sqrt{2}+2}{y+\sqrt{2}} \quad \text{dla } y\in(0,\sqrt{2}).$$ ■

Warunkowa wartość oczekiwana $E(Y|x)$ jest funkcją zmiennej x, $E(Y|x)=\varphi(x)$; funkcję tę nazywa się **regresją ZL Y względem ZL X**. Wykres tej funkcji na płaszczyźnie 0xy, czyli zbiór punktów postaci (x,y), gdzie $y=E(Y|x)$, nazywa się **linią regresji pierwszego rodzaju ZL Y względem ZL X** (krócej: linią regresji Y|X). W przypadku dyskretnym ta „linia" regresji jest zbiorem izolowanych punktów postaci $(x_i,E(Y|x_i))$.

Analogicznie definiuje się **regresję** oraz **linię regresji pierwszego rodzaju ZL X względem ZL Y** (linię regresji X|Y).

PRZYKŁAD 7.8. a) Dla danych z przykładu 7.7.a) linia regresji Y|X jest zbiorem dwóch punktów $\{(0,\frac{1}{2}),(1,\frac{2}{3})\}$. Podobnie zbiór $\{(\frac{1}{2},0),(\frac{2}{3},1)\}$ jest linią regresji X|Y.

b) Dla danych z przykładu 7.7.b) linia regresji Y|X jest odcinkiem prostej $y=\frac{1}{3}x$ dla $x\in(0,\sqrt{2})$. Natomiast łuk wykresu funkcji $x=2(y^2+y\sqrt{2}+2)/3(y+\sqrt{2})$ dla $y\in(0,\sqrt{2})$ jest linią regresji X|Y. ■

Zajmiemy się teraz przydatnością linii regresji. Załóżmy, że interesują nas wartości ZL Y, ale dostępne są nam tylko obserwacje wartości ZL X. Jeśli istnieje zależność między ZL X i Y, to pozostaje nam jedynie dobrać funkcję φ i stosować przybliżenie $Y\approx\varphi(X)$ oraz przy dobieraniu funkcji φ zadbać o minimalizację błędu tego przybliżenia. Przyjmujemy, że funkcja φ ma być tak dobrana, aby wyrażenie $E[Y-\varphi(X)]^2$ osiągało minimum. Dowodzi się, że wyrażenie to osiąga minimum, gdy jako funkcję φ obierzemy warunkową wartość oczekiwaną $E(Y|x)$, tj gdy $\varphi(x)=E(Y|x)$. Oznacza to, że (z pr-stwem 1) dla każdej funkcji φ spełniona jest nierówność:

$$E[Y-E(Y|X)]^2\le E[Y-\varphi(X)]^2. \tag{7.17}$$

Na przykład dla danych z przykładu 7.7.b) mamy (z pr-stwem 1):

$$E[Y-\frac{1}{3}X]^2\le E[Y-\varphi(X)]^2$$

przy dowolnej funkcji φ.

NIEZALEŻNOŚĆ ZMIENNYCH LOSOWYCH. Przypomnijmy: mówimy, że **ZL X, Y są niezależne**, gdy dla dowolnych $x\in R$, $y\in R$ spełniona jest równość

$$P(X<x,Y<y)=P(X<x)\cdot P(Y<y), \tag{7.18}$$

tj, gdy zdarzenia $\{\omega: X(\omega)<x\}$ i $\{\omega: Y(\omega)<y\}$ są niezależne. Gdy warunek ten nie jest spełniony, to ZL X, Y nazywa się **zależnymi**.

TWIERDZENIE 7.4 (trzy warunki konieczne i wystarczające niezależności ZL).

1) ZL X i Y są niezależne wtedy i tylko wtedy, gdy dla dowolnych zbiorów $A\subset R$ i $B\subset R$ zdarzenia $\{\omega: X(\omega)\in A\}$, $\{\omega: Y(\omega)\in B\}$ są niezależne, czyli

$$P(X\in A,Y\in B)=P(X\in A)\cdot P(Y\in B).$$

2) Niech (X,Y) oznacza WLS o punktach skokowych (x_i,y_j) i skokach p_{ij}. Wówczas ZL X i Y są niezależne wtedy i tylko wtedy, gdy dla każdego punktu skokowego (x_i,y_j)

$$P(X=x_i,Y=y_j)=P(X=x_i)\cdot P(Y=y_j) \quad \text{czyli} \quad p_{ij}=p_{i\cdot}\cdot p_{\cdot j}. \tag{7.19}$$

3) Niech (X,Y) oznacza WLC o GP f i brzegowych GP f_X i f_Y. Wówczas ZL X, Y są niezależne wtedy i tylko wtedy, gdy GP f WLC (X,Y) jest postaci:

$$f(x,y)=f_X(x)\cdot f_Y(y). \tag{7.20}$$

TWIERDZENIE 7.5. Jeżeli ZL X i Y są niezależne, to niezależne są również ZL $U=g(X)$ i $Z=h(Y)$, czyli: funkcje niezależnych zmien-nych losowych są niezależnymi zmiennymi losowymi.

PRZYKŁAD 7.9. ZL X i Y z przykładu 7.1 są zależne, bo np. $P(X=0,Y=0)=0,2\neq 0,4\cdot 0,4=P(X=0)\cdot P(Y=0)$. Zauważmy, że punkty skokowe $\{(0,0),(0,1),(1,0),(1,1)\}$ nie należą do wykresów funkcji postaci $y=g(x)$ lub $x=h(x)$, bo np. argumentowi 0 należałoby przyporządkować dwie różne wartości funkcji: $g(0)=0$ i $g(0)=1$. Zatem: *zależność (dokładnie: stochastyczna zależność) ZL X i Y nie oznacza ich zależności funkcyjnej*, tj nie oznacza istnienia funkcji g i h takich, że $Y=g(X)$ lub $X=h(Y)$. ■

PRZYKŁAD 7.10. ZL X i Y z przykładu 7.4 są zależne, ponieważ

$$f_X(x)\cdot f_Y(y)=x^3\cdot y(2-y^2)\neq 2xy=f(x,y).$$ ■

MOMENTY. Przyjmujemy, wykorzystywaną niżej, następującą terminologię.

1) Niech (X,Y) będzie WLS (o punktach skokowych (x_i,y_j) i skokach p_{ij}) albo WLC (o GP f). Niech poza tym będzie dana ciągła funkcja rzeczywista g zmiennych rzeczywistych x i y, $z=g(x,y)$. Funkcja g(X,Y) ZL X i Y jest pewną ZL, powiedzmy $g(X,Y)=U$.

Wartością oczekiwaną funkcji g(X,Y) **ZL** X i Y nazywa się liczbę Eg(X,Y) określoną wzorem:

$$\text{(7.21)}\quad Eg(X,Y)\overset{\text{def}}{=}\begin{cases}\sum\limits_{(x_i,y_j)} g(x_i,y_j)p_{ij}, & \text{gdy (X,Y) jest WLS,}\\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} g(x,y)f(x,y)dxdy, & \text{gdy (X,Y) jest WLC,}\end{cases}$$

przy założeniu, że szereg (gdy jest to "suma nieskończona") i całka (gdy jest to całka niewłaściwa) są bezwzględnie zbieżne.

2) **Momentem zwykłym** m_{kl} **rzędu** k+l **WL** (X,Y) nazywa się wartość oczekiwaną funkcji $g(X,Y)=X^kY^l$:

$$\text{(7.22)}\quad m_{kl}\overset{\text{def}}{=}E(X^kY^l).$$

3) **Momentem centralnym** μ_{kl} **rzędu** k+l **WL** (X,Y), mającego momenty $EX=m_X$ i $EY=m_Y$, nazywa się wartość oczekiwaną funkcji $g(X,Y)=(X-m_X)^k(Y-m_Y)^l$:

$$\text{(7.23)}\quad \mu_{kl}\overset{\text{def}}{=}E[(X-m_X)^k(Y-m_Y)^l].$$

4) Zanotujemy jeszcze pewne przypadki szczególne momentów i związków między nimi:

$m_{00}=1$, $m_{10}=EX=m_X$, $m_{01}=EY=m_Y$

$\mu_{00}=1$ (jeśli istnieją m_{10} i m_{01}), $\mu_{10}=E(X-m_X)=0$, $\mu_{01}=0$,

$$\mu_{20}=E(X-m_X)^2=VarX,\qquad \mu_{02}=E(Y-m_Y)^2=VarY,$$

$$\mu_{20}=m_{20}-m_{10}^2,\qquad \mu_{02}=m_{02}-m_{01}^2.$$

Ostatnie dwa związki wynikają ze znanej zależności:

$$VarX=EX^2-(EX)^2.$$

KOWARIANCJA. Zajmiemy się teraz bardziej szczegółowo momentem centralnym rzędu l+l, czyli momentem μ_{11}.

Kowariancją cov(X,Y) **ZL** X i Y, mających wartości oczekiwane m_X i m_Y, nazywamy moment centralny μ_{11}:

$$\text{(7.24)}\quad cov(X,Y)\overset{\text{def}}{=}\mu_{11}=E[(X-m_X)(Y-m_Y)].$$

Wzór ten, gdy uwzględnimy typ WL, przyjmuje postać

(7.25)

$$cov(X,Y)\overset{\text{def}}{=}\begin{cases}\sum\limits_{(x_i,y_j)}(x_i-m_X)(y_j-m_Y)p_{ij}, & \text{gdy (X,Y) jest WLS,}\\ \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x-m_X)(y-m_Y)f(x,y)dxdy, & \text{gdy (X,Y) jest WLC,}\end{cases}$$

TWIERDZENIE 7.6. Kowariancja cov(X,Y) ma następujące własności (X, Y - zmienne losowe dowolnego typu, mające momenty rzędu drugiego, a, b, c - stałe).

C1. $cov(X,Y)=EXY-EX\cdot EY$, *czyli* $\mu_{11}=m_{11}-m_{10}m_{01}$,

C2. $cov(X,X)=VarX$,

C3. $cov(aX+c,bY+d)=ab\,cov(X,Y)$,

C4. $|cov(X,Y)|\le\sigma_X\sigma_Y$; równość $|cov(X,Y)|=\sigma_X\sigma_Y$ zachodzi wtedy i tylko wtedy, gdy istnieją stałe a, b, c takie, że $P(aX+bY+c=0)=1$, tj gdy z pr-stwem 1 ZL X i Y są związane zależnością liniową.

C5. $Var(X\pm Y)=VarX+VarY\pm 2cov(X,Y)$.

D o w ó d . Ograniczymy się do dowodu własności C1 i C3 ułatwiających obliczanie kowariancji.

Ad. C1. $cov(X,Y)=E[(X-EX)(Y-EY)]=$

$$= E[XY - XEY - YEX + (EX)(EY)] =$$
$$= E(XY) - (EX)(EY) - (EY)(EX) + (EX)(EY) = E(XY) - (EX)(EY).$$

Ad. C3. $\operatorname{cov}(aX+c, bY+d) =$
$$= E[((aX+c) - E(aX+c))\,((bY+d) - E(bY+d))] =$$
$$= E[(aX - aEX)(bY - bEY)] = abE[(X - EX)(y - EY)] =$$
$$= ab\operatorname{cov}(X,Y).$$ □

Co mierzy, co charakteryzuje kowariancja $\operatorname{cov}(X,Y)$? Celem odpowiedzi przeanalizujmy pierwszy wzór (7.25) określający $\operatorname{cov}(X,Y)$ w przypadku dyskretnym. Iloczyny $(x_i - EX)(y_j - EY)$ są dodatnie, jeśli obydwa czynniki mają ten sam znak, tj gdy równocześnie x_i i y_j są duże i równocześnie x_i i y_j są małe. Oczywiście na wartość kowariancji mają wpływ również pr-stwa p_{ij}. Zatem: 1) *gdy* $\operatorname{cov}(X,Y)$ *jest duża i dodatnia (w sumie* (7.25) *składniki dodatnie "przeważyły" składniki ujemne), to ZL X i Y mają tendencję równoczesnego przyjmowania wartości dużych i równocześnie przyjmowania wartości małych oraz* 2) *gdy* $\operatorname{cov}(X,Y)$ *jest ujemna ale duża co do wartości bezwzględnej (w sumie* (7.25) *składniki ujemne przeważyły składniki dodatnie), to każda ze zmiennych X i Y ma tendencję przyjmowania wartości dużych, gdy druga przyjmuje wartości małe.*

Godny uwagi jest przypadek, gdy $\operatorname{cov}(X,Y) = 0$. Mówimy, że ZL X i Y są **nieskorelowane**, gdy ich kowariancja jest równa zero:

(7.26) $$\operatorname{cov}(X,Y) = 0.$$

Dla nieskorelowanych ZL z reguły nie występują tendencje, o których mówiliśmy wyżej.

Gdy $\operatorname{cov}(X,Y) \neq 0$, to mówimy, że ZL X i Y są **skorelowane**; przy tym, jeśli $\operatorname{cov}(X,Y) > 0$ ($\operatorname{cov}(X,Y) < 0$), to mówimy, że ZL X i Y są **skorelowane dodatnio** (**ujemnie**).

TWIERDZENIE 7.7. Niech ZL X i Y mają momenty rzędu drugiego. Wówczas: jeśli ZL X i Y są niezależne, to są one nieskorelowane.

D o w ó d . Zgodnie z własnością C1 kowariancji otrzymujemy:
$$\operatorname{cov}(X,Y) = E(XY) - EX \cdot EY.$$
Dla niezależnych ZL $E(XY) = (EX)\cdot(EY)$. Zatem $\operatorname{cov}(X,Y) = 0$ i tym samym ZL X, Y są nieskorelowane. □

Twierdzenie odwrotne do twierdzenia 7.7 jest fałszywe: nieskorelowane ZL mogą być (stochastycznie) zależne a nawet funkcyjnie zależne.

P R Z Y K Ł A D 7.11. ZLC X ma GP postaci:

$$f(x) = \begin{cases} 1-|x| & \text{dla } -1 < x < 1 \\ 0 & \text{dla } x \le -1 \vee x \ge 1. \end{cases}$$

Niech $Y = X^2$. Obliczymy $\operatorname{cov}(X,Y)$ korzystając z własności C1. Należy więc obliczyć EX, EY, E(XY):

$$EX = \int_{-1}^{1} x(1-|x|)dx = 0, \quad E(XY) = E(X\cdot X^2) = EX^3 = \int_{-1}^{1} x^3(1-|x|)dx = 0.$$

Zatem
$$\operatorname{cov}(X,Y) = 0 - 0\cdot EY = 0.$$
ZL X, Y są więc tu *nieskorelowane mimo, że są funkcyjnie zależne.*

WSPÓŁCZYNNIK KORELACJI.

Mimo, że $|\operatorname{cov}(X,Y)| \le \sigma_x\sigma_y$, to jednak nie istnieje stała C taka, że dla każdej pary ZL X, Y $|\operatorname{cov}(X,Y)| \le C$. Jest to pewną wadą kowariancji, ponieważ utrudnia ocenę ze względu na wielkość konkretnej obliczonej kowariancji. Wady tej nie ma inna charakterystyka pary ZL X, Y - charakteryzująca je "pod tym samym kątem" co kowariancja - współczynnik korelacji.

Współczynnikiem korelacji ZL X, Y, mających różne od zera odchylenie standardowe σ_X i σ_Y, nazywamy liczbę ρ_{XY} (krócej: ρ) określoną wzorem:

(7.27) $$\rho_{XY} \overset{def}{=} \frac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}.$$

Jeśli $\sigma_X \neq 0$ i $\sigma_Y \neq 0$, to:
1) ZL X, Y są nieskorelowane wtedy i tylko wtedy, gdy $\rho_{X,Y} = 0$,
2) ZL X,Y są skorelowane dodatnio wtedy i tylko wtedy, gdy $\rho_{X,Y} > 0$,
3) ZL X,Y są skorelowane ujemnie wtedy i tylko wtedy, gdy $\rho_{X,Y} < 0$.

Bezpośrednio z definicji (7.27) i własności kowariancji wynika prawdziwość następującego twierdzenia.

TWIERDZENIE 7.8. Współczynnik korelacji ρ_{XY} ma następujące własności (a, b, c, d - stałe, $a\neq 0$, $b\neq 0$)

ρ1. $\rho_{XY}=\dfrac{E(XY)-EX\cdot EY}{\sigma_X\sigma_Y}$,

ρ2. $\rho_{aX+c,bY+d}=\rho_{XY}\cdot\operatorname{sgn}(ab)$, gdzie $\operatorname{sgn}x=\begin{cases}-1, & gdy\ x<0\\ 0, & gdy\ x=0\\ 1, & gdy\ x>0,\end{cases}$

ρ3. $|\rho_{X,Y}|\le 1$, przy tym $|\rho_{X,Y}|=1$ wtedy i tylko wtedy gdy z pr-stwem 1 ZL X, Y są związane zależnością liniową: $P(aX+bY+c=0)=1$.

Twierdzenie przeciwstawne do twierdzenia 7.7 orzeka: jeśli $\rho_{X,Y}\neq 0$, to ZL X, Y są (stochastycznie) zależne. W szczególności, gdy $|\rho_{X,Y}|$ osiąga wartość maksymalną $|\rho_{X,Y}|=1$, to X, Y są, zgodnie z własnością ρ3, z pr-stwem 1 związane zależnością liniową. *Jeśli $|\rho_{X,Y}|$ mało różni się od jedności, to wartości (x,y) WL (X,Y) są (z pr-stwem 1) bardzo skupione wokół pewnej p r o s t e j*. Zatem współczynnik korelacji ρ_{XY} można interpretować jako *miarę stopnia z a l e ż n o ś c i l i n i o w e j* ZL X i Y.

P R Z Y K Ł A D 7.12. Doświadczenie polega na rzucie dwiema rzetelnymi kostkami sześciennymi i obserwacji liczby "wyrzuconych" na nich oczek. Określamy ZL: X - liczba wyrzuconych szóstek, Y - liczba wyrzuconych parzystych oczek. Wyznaczymy: a) funkcję pr-stwa WL (X,Y), b) cov(X,Y), c) Var(X+Y), Var(X−Y), d) ρ_{XY}.

a) Budujemy PP $(\Omega,\mathcal{A},P)$. Przestrzenią zdarzeń elementarnych jest zbiór $\Omega=\{(i,j): i,j=1,2,\ldots,6\}$, $\overline{\Omega}=6^2=36$; zbiorem zdarzeń jest rodzina $\mathcal{A}$ wszystkich podzbiorów zbioru Ω; pr-stwem P - pr-stwo zgodne z definicją klasyczną. Funkcja pr-stwa WL (X,Y) jest opisana następującą tabelą (bez ostatniego wiersza i ostatniej kolumny).

y_j \ x_i	0	1	2	$p_{\cdot j}$
0	9/36	0	0	9/36
1	12/36	6/36	0	18/36
2	4/36	4/36	1/36	9/36
$p_{i\cdot}$	25/36	10/36	1/36	1

Pokażemy sposób obliczania tylko części pr-stw p_{ij}:

$p_{00}=P(X=0,Y=0\})=P(\{11, 13, 15, 31, 33, 35, 51, 53, 55\})=9/36$,

$p_{11}=P(X=1,Y=1)=P(\{16, 36, 56, 65, 63, 61\})=6/36$,

$p_{12}=P(X=1,Y=2)=P(\{26, 46, 64, 62\})=4/36$.

b) $\operatorname{cov}(X,Y)=\{$ własność C1$\}=E(XY)-EX\cdot EY$,

$EX=\sum_{x_i}x_ip_{i\cdot}=12/36=1/3$, $\quad EY=\sum_{y_j}y_jp_{\cdot j}=1$,

$E(XY)=\sum_{(x_i,y_j)}x_iy_jp_{ij}=1\cdot 1\cdot\frac{6}{36}+1\cdot 2\cdot\frac{4}{36}+2\cdot 2\cdot\frac{1}{36}=\frac{18}{36}=\frac{1}{2}$,

$\operatorname{cov}(X,Y)=\frac{1}{2}-\frac{1}{3}\cdot 1=1/6$.

c) $EX^2=\sum_{x_i}x_i^2p_{i\cdot}=14/36=7/18$, $\operatorname{Var}X=EX^2-(EX)^2=\frac{7}{18}-\frac{1}{9}=\frac{5}{18}$,

$EY^2=\sum_{y_j}y_j^2p_{\cdot j}=\frac{54}{36}=\frac{3}{2}$, $\operatorname{Var}Y=EY^2-(EY)^2=\frac{3}{2}-1^2=\frac{1}{2}$;

$\operatorname{Var}(X+Y)=\operatorname{Var}X+\operatorname{Var}Y+2\operatorname{cov}(X,Y)=\frac{5}{18}+\frac{1}{2}+2\cdot\frac{1}{6}=\frac{10}{9}$,

$\operatorname{Var}(X-Y)=\operatorname{Var}X+\operatorname{Var}Y-2\operatorname{cov}(X,Y)=\frac{5}{18}+\frac{1}{2}-2\cdot\frac{1}{6}=\frac{4}{9}$.

d) $\rho_{X,Y}=\dfrac{\operatorname{cov}(X,Y)}{\sigma_x\sigma_y}=\frac{1}{6}\cdot\frac{3\sqrt{2}}{\sqrt{5}}\cdot\sqrt{2}=\frac{1}{\sqrt{5}}\approx 0{,}45$. ■

P R Z Y K Ł A D 7.13. Dla WL (X,Y) z przykładu 7.4 wyznaczymy: a) cov(X,Y), b) $\rho_{X,Y}$.

a) Korzystając z wcześniejszych obliczeń, otrzymujemy:

$$EX=\int_0^{\sqrt{2}}x\cdot x^3dx=\frac{4\sqrt{2}}{5}\approx 1{,}13,$$

$$EY=\int_0^{\sqrt{2}}y\cdot y(2-y^2)dy=\frac{8\sqrt{2}}{15}\approx 0{,}75,$$

$$E(XY)=\int_0^{\sqrt{2}}\Big[\int_0^{x}xy\cdot 2xydy\Big]dx=\frac{8}{9}\approx 0{,}89,$$

$$\operatorname{cov}(X,Y)=E(XY)-EX\cdot EY=\frac{8}{9}-\frac{4\sqrt{2}}{5}\cdot\frac{8\sqrt{2}}{15}=\frac{8}{225}\approx 0{,}04.$$

b) $EX^2=\int_0^{\sqrt{2}}x^2\cdot x^3dx=\frac{4}{3}$, $\operatorname{Var}X=\frac{4}{3}-(\frac{4\sqrt{2}}{5})^2=\frac{4}{3\cdot 25}$,

$EY^2=\int_0^{\sqrt{2}}y^2\cdot y(2-y^2)dy=\frac{2}{3}$, $\operatorname{Var}Y=\frac{2}{3}-(\frac{8\sqrt{2}}{15})^2=\frac{22}{225}$,

$$\rho_{X,Y}=\frac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}=\frac{8}{225}\cdot\frac{5\sqrt{3}}{2}\cdot\frac{15}{\sqrt{22}}=\frac{4}{\sqrt{66}}\approx 0{,}49.$$ ■

PROSTE REGRESJI DRUGIEGO RODZAJU. Rozważamy WL (X,Y). Podobnie jak w omówionym już temacie linii regresji pierwszego rodzaju, chcemy aproksymować (przybliżać) wartości ZL Y za pomocą wartości ZL X przyjmując $Y\approx\varphi(X)$. Jednak tym razem optymalnej funkcji φ, tj. funkcji φ minimalizującej wyrażenie $E[Y-\varphi(X)]^2$, poszukujemy nie wśród w s z y s t k i c h funkcji, lecz tylko wśród funkcji l i n i o w y c h: funkcja φ ma być postaci $\varphi(x)=\alpha x+\beta$.

Prostą

$$y=\alpha x+\beta$$

o tak dobranych parametrach α i β, aby wyrażenie

$$E[Y-(\alpha X+\beta)]^2,$$

traktowane jako funkcja parametrów α i β, osiągało wartość najmniejszą, nazywa się **prostą regresji drugiego rodzaju ZL Y względem ZL X** (krócej: prostą regresji Y|X).

TWIERDZENIE 7.9. Jeśli istnieją $\operatorname{cov}(X,Y)=\mu_{11}$ i $\operatorname{Var}X=\sigma_x^2\neq 0$, to prosta regresji drugiego rodzaju ZL Y względem ZL X (krócej: Y|X) jest postaci:

(7.28) $$y-m_Y=\frac{\mu_{11}}{\sigma_x^2}(x-m_x),\quad \text{gdzie } m_x=EX,\ m_Y=EY.$$

Analogicznie definiuje się **prostą regresji drugiego rodzaju ZL X względem ZL Y** (krócej: prostą regresji X|Y). Jej równanie - analogiczne do równania (7.28) - jest postaci:

(7.29) $$x-m_X=\frac{\mu_{11}}{\sigma_y^2}(y-m_y).$$

P R Z Y K Ł A D 7.14. Wyznaczymy równania obu prostych regresji drugiego rzędu dla danych z przykładu 7.13 korzystając z wykonanych tam już obliczeń.

Prosta regresji Y|X, zgodnie z (7.28), jest postaci:

$$y-1=\frac{1/6}{5/18}(x-\frac{1}{3}),\quad \text{czyli}\quad y=\frac{3}{5}x+\frac{4}{5}.$$

Podobnie prosta regresji X|Y, zgodnie z (7.29), ma równanie:

$$x-\frac{1}{3}=\frac{1/6}{1/2}(y-1),\quad \text{czyli}\quad x=\frac{1}{3}y.$$ ■

ZADANIA DO ROZWIĄZANIA.

1. Dana jest funkcja pr-stwa wektora losowego (X,Y):

y_j \ x_i	-10	0	10
0	0,10	0,15	0,20
10	0,15	0,25	0,15

a) Obliczyć pr-stwa: $p_1=P(X<10,Y<10)$, $p_2=P(X<10,Y\le 10)$, $p_3=P(X<10,Y>0)$, $p_4=P(X<10)$, $p_5=(X>-10)$, $p_6=P(-10\le X<10,0\le Y<10)$, $p_7=P(0\le X<20,-20\le Y<10)$, $p_8=P(X=0,Y=0)$, $p_9=P(X=0,Y=5)$.

b) Wyznaczyć brzegowe funkcje pr-stwa $p_{i\cdot}$ i $p_{\cdot j}$ oraz brzegowe dystrybuanty.

c) Obliczyć wartość oczekiwaną i wariancję ZL X i ZL Y.

d) Rozstrzygnąć, czy ZL X, Y są niezależne.

e) Wyznaczyć warunkową funkcję pr-stwa, warunkową dystrybuantę, warunkową wartość oczekiwaną i warunkową wariancję ZL X przy warunku, że ZL Y przyjęła wartość zero.

f) Wyznaczyć kowariancję i współczynnik korelacji ZL X, Y.

g) Wyznaczyć obie linie regresji pierwszego rodzaju.

h) Wyznaczyć obie proste regresji drugiego rodzaju.

2. Dana jest GP WL (X,Y): $f(x,y)=\begin{cases}8xy & \text{dla } 0<y\le x<1\\ 0 & \text{dla pozost.}(x,y)\in R^2.\end{cases}$

a) Obliczyć pr-stwa:

$p_1=P(X<\frac{1}{2},Y<\frac{1}{2})$, $p_2=P(X>\frac{1}{2},Y<1)$, $p_3=P(X>\frac{1}{2},Y>\frac{1}{2})$,

$p_4=P(X<\frac{1}{2})$, $p_5=P(Y<\frac{1}{2})$, $p_6=P(0<X<1,Y=\frac{1}{2})$.

b) Wyznaczyć brzegowe gęstości pr-stwa i brzegowe dystrybuanty.

c) Obliczyć wartość oczekiwaną i wariancję ZL X i ZL Y.

d) Rozstrzygnąć, czy ZL X, Y są niezależne.

e) Wyznaczyć warunkowe GP, warunkowe dystrybuanty, warunkowe wartości oczekiwane.

f) Wyznaczyć kowariancję i współczynnik korelacji ZL X, Y.

g) Wyznaczyć linie regresji pierwszego rodzaju.

h) Wyznaczyć obie proste regresji drugiego rodzaju.

3. Wyznaczyć gęstość pr-stwa wektora losowego (X,Y) mając dane: brzegową GP f_X i warunkową dystrybuantę $F_Y(\cdot|x)$

$f_X(x)=\begin{cases}e^{-x} & \text{dla } x>0\\ 0 & \text{dla } x\le 0,\end{cases}$ $F_Y(y|x)=\begin{cases}1-e^{-xy} & \text{dla } y>0\\ 0 & \text{dla } y\le 0,\end{cases}$ $x>0$.

4. Rozstrzygnąć, czy ZL X i Y są: 1) niezależne, 2) nieskorelowane, gdy wektor losowy ciągły (X,Y) ma:

a) gęstość pr-stwa $f(x,y)=\begin{cases}e^{-x-y} & \text{dla } x>0\wedge y>0\\ 0 & \text{dla } x\le 0\vee y\le 0,\end{cases}$

b) gęstość pr-stwa $f(x,y)=\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}$ dla $(x,y)\in R^2$,

c) rozkład jednostajny w kwadracie D o wierzchołkach (-1,0), (0,-1), (1,0), (0,1), tj rozkład zadany gęstością pr-stwa (2 jest polem danego kwadratu D):

$$f(x,y)=\begin{cases}\frac{1}{2} & \text{dla } (x,y)\in D\\ 0 & \text{dla pozost. } (x,y),\end{cases}$$

d) rozkład jednostajny w kole $x^2+y^2\le r^2$.

5. Oszacować z góry cov(X,Y), jeśli VarX=25, VarY=4.

6. Wykazać, że Var(XY)=VarX·VarY wtedy i tylko wtedy, gdy EX=0 i EY=0.

7. Obliczyć Var(XY) wiedząc, że ZL X, Y są niezależne oraz EX=−2, EY=3, VarX=3, VarY=1.

8. Obliczyć Var(3X−4Y), jeśli VarX=1, VarY=2, $\operatorname{cov}(X,Y)=\frac{1}{2}$.

9. a) Obliczyć cov(U,V), jeśli: U=2X−3, V=−3X+2, VarX=2.

b) Obliczyć $\rho_{X,2X+1}$, $\rho_{X,-2X+1}$, $\rho_{X-3,-3X+2}$.

c) Dokonujemy 5 doświadczeń Bernoulliego z pr-stwem p sukcesu w pojedynczym doświadczeniu. Niech X oznacza liczbę sukcesów w tych pięciu doświadczeniach, Y - liczbę porażek. Obliczyć: cov(X,Y), $\rho_{X,Y}$.

10. a) ZL X przyjmuje wartości −3, 0, 3 odpowiednio z pr-stwami: 0,25, 0,5, 0,25. Niech $Y=X^2$. Wykazać, że ZL X,Y są nieskorelowane (mimo, że są funkcyjnie zależne).

b) ZL X ma gęstość pr-stwa postaci:

$$f(x)=\begin{cases}\frac{1}{2}\cos x & \text{dla } -\frac{\pi}{2}<x<\frac{\pi}{2}\\ 0 & \text{dla pozost. } x.\end{cases}$$

Niech Y=|X|. Sprawdzić, że ZL X, Y są nieskorelowane.

11. a) Jaki punkt jest punktem wspólnym obu prostych regresji drugiego rodzaju?

b) Kiedy proste regresji drugiego rodzaju pokrywają się?

c) Kiedy proste regresji drugiego rodzaju są prostopadłe do osi układu?

d) Prosta y=2x−3 jest linią regresji pierwszego rodzaju ZL Y|X. Napisać równanie prostej regresji drugiego rodzaju ZL Y|X. Wynik uzasadnić.

Odpowiedzi.

1. a) $p_1=0{,}25$, $p_2=0{,}65$, $p_3=0{,}40$, $p_4=0{,}65$, $p_5=0{,}75$, $p_6=0{,}25$, $p_7=0{,}35$, $p_8=0{,}15$, $p_9=0$.

b)

x_i	−10	0	10
p_i	0,25	0,40	0,35

y_j	0	10
p_j	0,45	0,55

c) $EX=1$, $EY=5{,}5$, $VarX=59$, $VarY=24{,}75$;

d) zależne;

e)

x_i	-10	0	10
$P(X=x_i\|Y=0)$	2/9	3/9	4/9

x	$(-\infty,-10>$	$(-10,0>$	$(0,10>$	$(10,\infty)$
$F_X(x\|Y=0)$	0	2/9	5/9	1

$E(X|Y=0)=20/9$, $Var(X|Y=0)=5000/81\approx 61{,}7$;

f) $cov(X,Y)=-5{,}5$, $\rho_{X,Y}\approx -0{,}14$;

g)

y	0	10
E(X|y)	20/9	0

x	-10	0	10
E(Y|x)	20/9	25/4	30/7

h) $y-5{,}5=\frac{-5{,}5}{59}(x-1)$, $y-5{,}5=\frac{24{,}75}{-5{,}5}(x-5{,}5)$.

2. a) $p_1=1/16$, $p_2=15/16$, $p_3=9/16$, $p_4=1/16$, $p_5=7/16$, $p_6=0$;

b) $f_X(x)=\begin{cases}4x^3 & 0<x<1\\ 0 & \text{dla } x\le 0 \vee x\ge 1\end{cases}$ $f_Y(y)=\begin{cases}4y(1-y^2) & \text{dla } 0<y<1\\ 0 & \text{dla } y\le 0 \vee y\ge 1\end{cases}$;

c) $EX=4/5$, $EY=8/15$, $VarX=2/75$, $VarY=11/225$; d) zależne;

e) Dla $y\in(0,1)$ $f_X(x|y)=\begin{cases}2x/(1-y^2) & \text{dla } y<x<1\\ 0 & \text{dla } x\le y \vee x\ge 1\end{cases}$

Dla $x\in(0,1)$ $f_Y(y|x)=\begin{cases}2y/x^2 & \text{dla } 0<y<x\\ 0 & \text{dla } y\le 0 \vee y\ge x\end{cases}$.

$E(X|y)=\frac{2(y^2+y+1)}{3(y+1)}$ dla $y\in(0,1)$; $E(Y|x)=\frac{2}{3}x$ dla $0<x<1$;

f) $cov(X,Y)=4/225$, $\rho_{X,Y}=4/\sqrt{66}\approx 0{,}49$; g) jak w punkcie e);

h) $y-\frac{8}{15}=\frac{2}{3}(x-\frac{4}{5})$, $y-\frac{8}{15}=\frac{11}{4}(x-\frac{4}{5})$.

3. $f(x,y)=xe^{-x(y+1)}$, $x>0$, $y>0$. 4. a) niezależne (i nieskorelowane), b) niezależne, c) zależne i nieskorelowane, d) zależne i nieskorelowane.

5. $cov(X,Y)\le 10$; 7. 34; 8. 29; 9. a) -12, b) 1, -1, -1; c) -1;

11. a) (EX,EY), b) $|\rho|=1$, c) $|\rho|=0$, d) $y=2x-3$.

VI. ELEMENTY STATYSTYKI MATEMATYCZNEJ

1. TRZY ROZKŁADY WAŻNE W STATYSTYCE.

WARIANCJA ŚREDNIEJ ARYTMETYCZNEJ. Wiedzę o wariancji uzupełnimy teraz o wariancję sumy n zmiennych losowych (ZL) i ich średniej arytmetycznej (stosujemy skróty z rozdziału V).

TWIERDZENIE 1.1. Jeżeli ZL $X_1,X_2,\dots,X_n$ są niezależne i mają jednakowe wartości oczekiwane i jednakowe wariancje:

$$EX_1=EX_2=\dots=EX_n=\mu,\qquad VarX_1=VarX_2=\dots=VarX_n=\sigma^2$$

(w szczególności ma to miejsce, gdy ZL X_i mają ten sam rozkład pr-stwa), to:

1) wartość oczekiwana i wariancja ich sumy

$$S_n=X_1+X_2+\dots+X_n\equiv\sum_{i=1}^{n}X_i$$

jest n-krotnością odpowiednio wartości oczekiwanej i wariancji składników X_i:

$$ES_n=n\cdot EX_i=n\mu,\qquad VarS_n=n\cdot VarX_i=n\sigma^2, \tag{1.1}$$

2) ich średnia arytmetyczna

$$\overline{X}_n\overset{def}{=}\frac{1}{n}(X_1+X_2+\dots+X_n)\equiv\frac{1}{n}\sum_{i=1}^{n}X_i\equiv\frac{1}{n}S_n$$

ma tę samą wartość oczekiwaną co składniki X_i i wariancję n razy mniejszą (odchylenie standardowe $\sqrt{n}$ razy mniejsze):

$$(1.2)\qquad E\overline{X}_n = EX_i = \mu, \quad Var\overline{X}_n = \frac{VarX_i}{n} = \frac{\sigma^2}{n}, \quad D\overline{X} = \frac{\sigma}{\sqrt{n}}.$$

D o w ó d . Korzystając z odpowiednich własności wartości oczekiwanej i wariancji, otrzymujemy:

$$ES_n = E(X_1+X_2+\cdots+X_n) = EX_1+EX_2+\cdots+EX_n = \{ EX_i = \mu \} = n\mu,$$

$$VarS_n = Var(X_1+X_2+\cdots+X_n) = \{ \text{ZL } X_1,X_2,\ldots,X_n \text{ są niezależne} \} =$$

$$= VarX_1+VarX_2+\cdots+VarX_n = \{ VarX_i = \sigma^2 \} = n\sigma^2,$$

$$E\overline{X}_n = E(S_n/n) = \frac{1}{n}ES_n = \frac{1}{n}n\mu = \mu,$$

$$Var\overline{X}_n = Var(S_n/n) = \frac{1}{n^2}VarS_n = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}.$$

□

Z ostatniego twierdzenia wynika, *że średnia arytmetyczna* $\overline{X}_n$ ZL $X_1,X_2,\ldots,X_n$ *jest znacznie bardziej skupiona wokół ich wspólnej wartości oczekiwanej* μ *niż składniki* X_i (por. rys 6.3 w rozdziale V wykonany dla $DX_i = 1$, $n = 4$). Stanowi to teoretyczne uzasadnienie powszechnego w praktyce postępowania: za wartość pomiaru określonej wielkości przyjmuje się nie wynik pojedynczego pomiaru, lecz średnią arytmetyczną kilku pomiarów tej wielkości.

ZMIENNA O ROZKŁADZIE NORMALNYM (ciąg dalszy). Znany już nam rozkład normalny jest najważniejszym rozkładem pr-stwa. Trudno byłoby przecenić jego rolę zarówno w samej teorii jak również w praktycznych zastosowaniach, w tym w statystyce matematycznej. Jego fundamentalne znaczenie wynika z centralnych twierdzeń granicznych rachunku pr-stwa (przykład takiego twierdzenia poznamy w następnym paragrafie) i stąd, że ZL wykorzystywane w statystyce matematycznej często są funkcjami ZL o rozkładzie normalnym (przekonamy się o tym już w tym paragrafie). Obecnie przytoczymy dwa dalsze twierdzenia o rozkładzie normalnym.

TWIERDZENIE 1.2. Jeżeli ZL $X_1,X_2,\ldots,X_n$ są niezależne i mają ten sam rozkład normalny $N(\mu,\sigma)$, to ich suma $S_n = X_1+X_2+\cdots+X_n$ ma rozkład normalny $N(n\mu,\sigma\sqrt{n})$.

Dowód faktu, że ZL S_n ma rozkład normalny wymaga znajomości szerszego zakresu rachunku pr-stwa. Natomiast równości $ES_n = n\mu$ i $VarS_n = n\sigma^2$ są prawdziwe dla dowolnego, wspólnego rozkładu składników X_i (por. tw. 1.1).

TWIERDZENIE 1.3. Jeżeli ZL $X_1,X_2,\ldots,X_n$ są niezależne i mają ten sam rozkład normalny $N(\mu,\sigma)$, to ich średnia arytmetyczna

$$\overline{X} = \frac{1}{n}(X_1+X_2+\cdots+X_n) \equiv \frac{1}{n}\sum_{i=1}^{n}X_i$$

ma rozkład normalny $N(\mu,\sigma/\sqrt{n})$.

D o w ó d . Równości $E\overline{X} = \mu$, $Var\overline{X} = \sigma^2/n$, czyli $DX = \sigma/\sqrt{n}$ znamy z twierdzenia 1.1. Natomiast normalność rozkładu średniej arytmetycznej wynika z następujących przesłanek: 1) S_n ma rozkład normalny, 2) funkcja liniowa $Y = aX+b$ ZL X o rozkładzie normalnym ma rozkład normalny (widzieliśmy to w dowodzie twierdzenia 6.1 z rozdziału V), 3) średnia arytmetyczna $\overline{X} = S_n/n$ jest funkcją liniową ZL S_n. □

P R Z Y K Ł A D 1.1. ZL $X_1,X_2,\ldots,X_{25}$ są niezależne i mają ten sam rozkład normalny N(3,4), S_{25} i $\overline{X}_{25}$ oznaczają ich sumę i średnią arytmetyczną. Obliczymy pr-stwa: a) $P(|S_{25}-75|<1)$, b) $P(|\overline{X}_{25}-3|<1)$.

a) $ES_{25} = 25\cdot 3 = 75$, $VarS_{25} = 25\cdot 4^2$, $DS_{25} = 5\cdot 4 = 20$, $S_{25} \sim N(75;20)$.

Zatem

$$P(|S_{25}-75|<1) = P(-1<S_{25}-75<1) = \Phi(\frac{1}{20})-\Phi(-\frac{1}{20}) =$$

$$= 2\Phi(0,05)-1 = 0,0388 \approx 0,04.$$

Pr-stwo, że suma S_{25} odchyli się co do modułu od swej wartości oczekiwanej o mniej niż 1, okazało się bardzo małe. Należy to tłumaczyć dużą wartością odchylenia standardowego $DS_{25} = 20$, a więc dużym rozrzutem wartości ZL S_{25} wokół jej wartości oczekiwanej $ES_{25} = 75$.

b) $E\overline{X}_{25} = 3$, $Var\overline{X}_{25} = 4^2/25$, $D\overline{X}_{25} = 2/5 = 0,4$, $\overline{X}_{25} \sim N(3;0,4)$.

Zatem:

$$P(|\overline{X}_{25}-3|<1) = P(-1<\overline{X}_{25}-3<1) = \Phi(\frac{1}{0,4})-\Phi(-\frac{1}{0,4}) = 2\Phi(2,5)-1 =$$

$$= 2\cdot 0,9938-1 = 0,9876 \approx 0,99.$$

Pr-stwo tego, że średnia arytmetyczna $\overline{X}_{25}$ odchyli się (co do modułu) od swej wartości oczekiwanej o mniej niż 1 okazało się tym razem bardzo duże. Jest to spowodowane małą wartością odchylenia standardowego $D\overline{X}_{25}=0{,}4$ i tym samym małym rozrzutem wartości ZL $\overline{X}_{25}$ względem jej wartości oczekiwanej $E\overline{X}_{25}=3$. ■

Omówimy obecnie dwa rozkłady pr-stwa związane z rozkładem normalnym. Mają one podstawowe znaczenie w statystyce matematycznej.

ZMIENNA O ROZKŁADZIE CHI-KWADRAT. Mówimy, że ZLC X ma **rozkład chi-kwadrat** z ν stopniami swobody, $\nu=1,2,\ldots$, jeśli jej GP γ_ν jest postaci

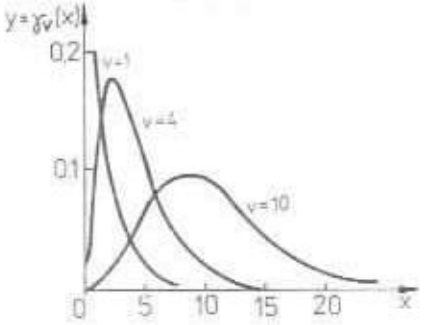
$$\gamma_\nu(x)=\begin{cases} C_\nu x^{\frac{1}{2}\nu-1}e^{-\frac{1}{2}x} & \text{dla } x>0,\\ 0 & \text{dla } x\le 0,\end{cases} \tag{1.3}$$

gdzie C_ν jest stałą zależną tylko od liczby stopni swobody ν i tak dobraną, aby spełniony był warunek unormowania.

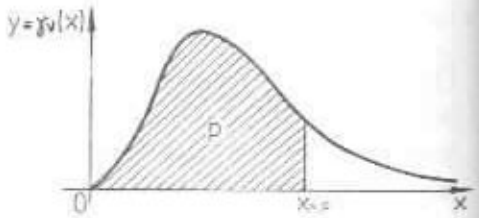
Oblicza się, że dla ZL X o tym rozkładzie wartość oczekiwana EX i Var X wyrażają się wzorami:

$$EX=\nu, \qquad VarX=2\nu. \tag{1.4}$$

Krzywe gęstości dla liczby stopni swobody $\nu=1$, $\nu=4$, $\nu=10$ są podane na rysunku 1.1.



Rys 1.1



Rys 1.2

Istnieją tablice (Tablica III na końcu książki) kwantyli ZL X o rozkładzie chi-kwadrat, tj liczb $x_{\nu,p}$ spełniających warunek (por. rys. 1.2):

$$\int_{-\infty}^{x_{\nu,p}}\gamma_\nu(x)dx=p.$$

PRZYKŁAD 1.2. ZL X ma rozkład chi-kwadrat z 7. stopniami swobody. Wyznaczymy liczbę x spełniającą warunek $P(X<x)=0{,}99$.

Szukaną liczbą jest kwantyl $x=x_{7;0,99}$. Z tablicy III odczytujemy $x_{7;0,99}=18{,}48$. ■

Związek rozkładu chi-kwadrat z rozkładem normalnym jest treścią następującego twierdzenia.

TWIERDZENIE 1.4. Jeżeli $X_1,X_2,\ldots,X_\nu$ są niezależnymi ZL o tym samym standaryzowanym rozkładzie normalnym N(0,1), to ZL

$$X \overset{def}{=} X_1^2+X_2^2+\cdots+X_\nu^2 \tag{1.5}$$

ma rozkład chi-kwadrat z ν stopniami swobody.

Z twierdzenia tego wynika również sens probabilistyczny liczby stopni swobody: liczba stopni swobody ν jest liczbą n i e z a l e ż n y c h składników w sumie (1.5).

Zastosowanie ZL o rozkładzie chi-kwadrat poznamy w ostatnim paragrafie tego rozdziału.

ZMIENNA O ROZKŁADZIE STUDENTA. Mówimy, że ZL t_k ma **rozkład Studenta** z k stopniami swobody, $k=1,2,\ldots$, gdy jej GP jest postaci:

$$g_k(x)=C_k\left(1+\frac{1}{k}x^2\right)^{-(k+1)/2}, \qquad -\infty<x<\infty, \tag{1.6}$$

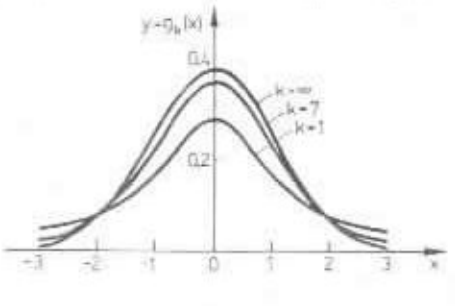
gdzie C_k jest stałą zależną tylko od liczby k stopni swobody i tak dobraną, aby spełniony był warunek unormowania.

Krzywa gęstości (1.6) jest symetryczna względem osi 0y (por. rys 1.3) i zależy od liczby stopni swobody; ze wzrostem k szybko upodabnia się do krzywej normalnej N(0,1), prawdziwe jest bowiem

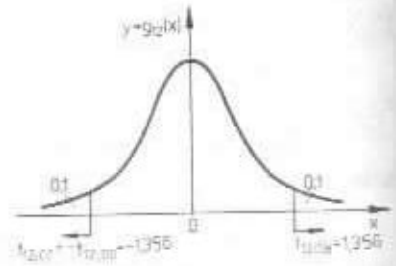
TWIERDZENIE 1.5. GP g_k rozkładu Studenta, gdy liczba stopni swobody $k\to\infty$, szybko dąży do GP standaryzowanego rozkładu normalnego N(0;1).

Z symetrii względem osi 0y krzywej gęstości Studenta wynika, że podobnie jak dla ZL U o rozkładzie N(0,1), kwantyle ZL t_k o rozkładzie Studenta spełniają warunek:

(1.7) $$t_{k,p} = -t_{k,1-p}.$$



Rys 1.3.



Rys 1.4.

P R Z Y K Ł A D 1.3. ZL X ma rozkład Studenta z k=12 stopniami swobody. Chcemy znaleźć liczbę x spełniającą warunek: $P(X<x)=0{,}1$.

Mamy: $P(X<x)=P(t_{12}<x)=0{,}1$. Szukaną liczbą jest więc kwantyl $t_{12;0,1}$. Tablica II kwantyli rozkładu Studenta zawiera tylko kwantyle rzędu $p\geq 0{,}9$. Pomocna jest tu równość (1.7), (rys. 1.4): $t_{12;0,1}=-t_{12;0,9}=-1{,}356$ ■

Wartość oczekiwana i wariancja ZL t_k o rozkładzie Studenta z k stopniami swobody wyrażają się wzorami (Et_k istnieje dla $k>1$, $Var\,t_k$ istnieje dla $k>2$).

(1.8) $$Et_k = 0, \qquad Var\,t_k = k/(k-2).$$

Związek rozkładu Studenta z rozkładem normalnym i sens liczby stopni wynika z następującego twierdzenia.

TWIERDZENIE 1.6. Jeżeli: 1) ZL U i Y są niezależne, 2) ZL U ma rozkład normalny N(0;1), 3) ZL Y ma rozkład chi-kwadrat z k stopniami swobody, to ZL Z

$$Z = U/\sqrt{Y/k}$$

ma rozkład Studenta z k stopniami swobody.

Zastosowanie rozkładu Studenta poznamy w paragrafach 5 i 7.

ZADANIA DO ROZWIĄZANIA.

1. ZL $X_1,X_2,\ldots,X_{16}$ są niezależne i mają ten sam rozkład normalny N(1;2); S_{16} i $\overline{X}_{16}$ oznaczają ich sumę i średnią arytmetyczną. Obliczyć: a) $P(|S_{16}-16|<8)$, b) $P(|\overline{X}_{16}-1|<1)$.
2. ZL $X_1,X_2,\ldots,X_{10}$ są niezależne i mają ten sam rozkład N(0;1). Wyznaczyć liczbę a spełniającą warunek $P(X_1^2+X_2^2+\cdots+X_{10}^2<a)=0{,}95$.
3. ZL $X_1,X_2,\ldots,X_{11}$ są niezależne i mają ten sam rozkład normalny N(0;1). Wyznaczyć liczbę a spełniającą warunek

$$P\left(\frac{X_{11}}{\sqrt{(X_1^2+X_2^2+\cdots+X_{10}^2)/10}}<a\right)=0{,}95.$$

Odpowiedzi. 1. a) 0,68; b) 0,95. 2. a=18,31. 3. a=1,812.

2. PRZYKŁADY TWIERDZEŃ GRANICZNYCH.

W paragrafie 5 rozdziału V poznaliśmy twierdzenie o zbieżności ciągu rozkładów dwumianowych do rozkładu Poissona. Jest to przykład twierdzenia granicznego. Twierdzenia graniczne są ważnym działem rachunku pr-stwa i stanowią podstawę wielu procedur statystycznych. Obecnie sformułujemy najprostsze, klasyczne przykłady takich twierdzeń.

SŁABE PRAWA WIELKICH LICZB.

TWIERDZENIE 2.1 (słabe prawo wielkich liczb Chinczyna). Jeżeli $X_1,X_2,\ldots,X_n,\ldots$ jest ciągiem niezależnych ZL o jednakowym rozkładzie pr-stwa i skończonej wartości oczekiwanej $EX_i=\mu$, to przy dowolnym $\varepsilon>0$ i $n\to\infty$

(2.1) $$P(|\overline{X}_n-\mu|<\varepsilon)\to 1,$$

gdzie $\overline{X}_n$ oznacza średnią arytmetyczną ZL $X_1,X_2,\ldots,X_n$.

Prawo to orzeka, że pr-stwo odchylenia (co do modułu) średniej arytmetycznej $\overline{X}_n$ ZL $X_1,X_2,\ldots,X_n$ od ich wspólnej wartości oczekiwanej $\mu=EX_i$ o mniej niż dowolnie wzięta liczba $\varepsilon>0$, jest dowolnie bliskie jedności, gdy weźmie się dostatecznie dużą liczbę n ZL $X_1,X_2,\ldots,X_n$.

Jeśli zatem nie jest znana wspólna wartość oczekiwana $\mu=EX_i$ i dysponujemy wartościami $x_1,x_2,\ldots,x_n$ odpowiednio ZL $X_1,X_2,\ldots,X_n$, to przy dużym n możemy przyjąć oszacowanie $\mu\approx\overline{x}$, gdzie $\overline{x}$ jest wartością ZL $\overline{X}_n$, czyli

$$\overline{x}=\frac{1}{n}(x_1+x_2+\ldots+x_n).$$

Bezpośrednio z twierdzenia 2.1 - jeśli uwzględnić własność (5.3) z rozdziału V - wynika historycznie najstarsze

TWIERDZENIE 2.2 (słabe prawo wielkich liczb Bernoulliego). Jeżeli $S_1,S_2,\ldots,S_n,\ldots$ są ZL odpowiednio o rozkładach dwumianowych b(1,p), b(2,p), ... ,b(n,p), ... , gdzie $0<p<1$, to przy dowolnym $\varepsilon>0$

$$P(|\frac{S_n}{n}-p|<\varepsilon)\to 1, \quad \text{gdy } n\to\infty. \tag{2.2}$$

Zmienną losową $\overline{X}_n=S_n/n$ nazywa się **częstością** zdarzenia A w serii n doświadczeń Bernoulliego. Zatem pr-stwo tego, że - częstość S_n/n zdarzenia A w serii n doświadczeń Bernoulliego różni się (co do modułu) od pr-stwa $p=P(A)$ zajścia zdarzenia A w pojedynczym doświadczeniu dowolnie mało - jest dowolnie bliskie jedności, gdy n jest dostatecznie duże. Stąd wynika dalej, że: gdy n jest duże, p nie jest znane, to zaobserwowaną wartość s_n/n częstości względnej S_n/n możemy przyjąć za oszacowanie nieznanego pr-stwa $p=P(A)$ sukcesu A w pojedynczym doświadczeniu: $p\approx s_n/n$.

MOCNE PRAWA WIELKICH LICZB.

TWIERDZENIE 2.3 (mocne prawo wielkich liczb Kołmogorowa). Jeżeli $X_1,X_2,\ldots,X_n,\ldots$ jest ciągiem niezależnych ZL o jednakowym rozkładzie pr-stwa i skończonej wartości oczekiwanej $EX_i=\mu$, to

$$P(\lim_{n\to\infty}\overline{X}_n=\mu)=1. \tag{2.3}$$

Równość ta orzeka, że z *pr-stwem 1 ciąg średnich arytmetycznych* $(\overline{X}_n)$ *ZL* $X_1,X_2,\ldots,X_n$ *jest zbieżny do wspólnej wartości oczekiwanej* $\mu=E(X_i)=E(\overline{X}_n)$. Istnieje zatem zdarzenie A takie, że $P(A)=1$ oraz dla każdego zdarzenia elementarnego $\omega\in A$ ciąg liczbowy $(\overline{X}_n(\omega))$ jest zbieżny (w zwykłym sensie) do wspólnej wartości oczekiwanej μ.

Równość (2.3) jest ponownym uzasadnieniem zasadności przyjęcia oszacowania $\mu\approx\overline{x}$, gdy μ nie jest znane.

PRZYKŁAD 2.1. Załóżmy, że ZL $X_1,X_2,\ldots,X_n,\ldots$ są niezależne i mają ten sam rozkład geometryczny z parametrem p. Wiemy, że $EX_i=1/p$. Rozważmy średnie arytmetyczne $\overline{X}_n=(X_1+X_2+\ldots+X_n)/n$. Zgodnie z tezą mocnego prawa wielkich liczb Kołmogorowa, równość $\lim_{n\to\infty}\overline{X}_n=1/p$ spełniona jest z pr-stwem 1, tzn. zdarzenie

$$A=\{\omega:\lim_{n\to\infty}\overline{X}_n(\omega)=1/p\} \quad \text{ma pr-stwo 1, } P(A)=1. \qquad \blacksquare$$

Bezpośrednio z twierdzenia 2.3 i własności (5.3) z rozdziału V wynika następujące

TWIERDZENIE 2.4 (mocne prawo wielkich liczb Borela). Jeżeli $S_1,S_2,\ldots,S_n,\ldots$ są ZL odpowiednio o rozkładach dwumianowych b(1,p), b(2,p), ..., b(n,p), ..., to przy $n\to\infty$

$$P(\lim\frac{S_n}{n}=p)=1. \tag{2.4}$$

Równość ta orzeka, że z pr-stwem 1 częstość S_n/n zdarzenia A w serii n doświadczeń Bernoulliego jest przy $n\to\infty$ zbieżna do pr-stwa $p=P(A)$ zdarzenia A w pojedynczym doświadczeniu. Zatem równość (2.4) jest ponownym uzasadnieniem oszacowania $p\approx s_n/n$.

CENTRALNE TWIERDZENIE GRANICZNE.

Często spotyka się sytuacje, w których interesujący nas wynik doświadczenia jest rezultatem dużej liczby czynników, które działają niezależnie, których działania sumują się i takich, że wielkość oddziaływania każdego z nich jest nieznaczna w porównaniu z sumą oddziaływania pozostałych czynników. Taką sytuację można zaobserwować np. w mechaniźmie powstawania błędów losowych. W praktyce w takich sytuacjach zakłada się, że wynik doświadczenia jest realizacją ZL o rozkładzie normalnym.

Teoretycznego wyjaśnienia zasadności takiego postępowania dostarczają **centralne twierdzenia graniczne** (CTG) rachunku pr-stwa. Tak nazywa się twierdzenia, które orzekają o zbieżności rozkładu sum zmiennych losowych do rozkładu normalnego, gdy liczba składników dąży do nieskończoności. Zacytujmy jedno z najprostszych takich twierdzeń i przypadek szczególny tego twierdzenia.

TWIERDZENIE 2.5 (CTG Lindeberga-Levy'ego). Jeżeli ZL $X_1, X_2, \ldots, X_n, \ldots$ są niezależne, mają ten sam rozkład pr-stwa, skończoną wartość oczekiwaną $\mu = EX_i$, skończoną i różną od zera wariancję $\sigma^2 = VarX_i$, to ciąg dystrybuant (F_n) standaryzowanych ZL U_n

$$(2.5) \qquad U_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}, \quad \text{gdzie } S_n = X_1 + X_2 + \ldots + X_n$$

jest zbieżny na całej prostej do dystrybuanty Φ standaryzowanego rozkładu normalnego N(0,1):

$$(2.6) \qquad \lim_{n\to\infty} F_n(u) = \Phi(u), \quad \textit{czyli} \quad \lim_{n\to\infty} P(U_n < u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{1}{2}t^2}\, dt.$$

Z (2.6) wynika również, że

$$(2.7) \qquad \lim_{n\to\infty} P(u_1 < U_n < u_2) = \frac{1}{\sqrt{2\pi}} \int_{u_1}^{u_2} e^{-\frac{1}{2}t^2}\, dt = \Phi(u_2) - \Phi(u_1).$$

Dla dużych n z (2.6) i (2.7) wynikają następujące **przybliżenia normalne** dla standaryzowanych sum U_n:

$$(2.8) \qquad P(U_n < u) \overset{CTG}{\approx} \Phi(u), \quad P(u_1 < U_n < u_2) \overset{CTG}{\approx} \Phi(u_2) - \Phi(u_1).$$

PRZYKŁAD 2.2. Każdą ze 192 liczb z dużą ilością cyfr po przecinku zaokrąglamy do części całkowitej. Przyjmujemy, że błąd zaokrąglenia X_i i-tej liczby, $i = 1,2,\ldots,192$, jest ZL o rozkładzie prostokątnym $R(-0,5; 0,5)$. Interesujący nas błąd sumy tych 192 liczb jest wówczas ZL postaci $S_{192} = X_1 + X_2 + \ldots + X_{192}$. Obliczymy pr-stwo, że błąd ten co do modułu będzie mniejszy od 10: $P(|S_{192}| < 10)$.

Posłużymy się przybliżeniem (2.8). Korzystając ze wzorów na wartość oczekiwaną i wariancję ZL o rozkładzie prostokątnym R(a,b), otrzymujemy:

$$EX_i = \frac{1}{2}(a+b) = 0, \quad VarX_i = \frac{1}{12}(b-a)^2 = \frac{1}{12},$$

$$U_{192} = \frac{S_{192} - 192\cdot 0}{\frac{1}{2\sqrt{3}}\sqrt{192}} = \frac{S_{192}}{4}.$$

Zatem według (2.8) mamy:

$$P(|S_{192}| < 10) = P(-10 < S_{192} < 10) = P\left(\frac{-10}{4} < \frac{S_{192}}{4} < \frac{10}{4}\right) =$$

$$\approx \Phi(2,5) - \Phi(-2,5) = 2\Phi(2,5) - 1 = 2\cdot 0,9938 - 1 = 0,9876 \approx 0,99.$$

Szukane pr-stwo w przybliżeniu wynosi 0,99. Na zakończenie przypomnijmy, że zgodnie z teorią błędów, maksymalny bezwzględny błąd sumy jest równy sumie maksymalnych błędów bezwzględnych składników. W rozważanym przykładzie należałoby zatem przyjąć, że jest on równy $192\cdot 0,5 = 96$. ■

Z ostatniego twierdzenia - jeśli uwzględnić równość (5.3) z rozdziału V oraz równości $ES_n = np$, $VarS_n = npq$, $DS_n = \sqrt{npq}$ - wynika

TWIERDZENIE 2.6 (Moivre'a - Laplace'a). Jeżeli (S_n) jest ciągiem ZL o rozkładach dwumianowych b(n,p), to ciąg dystrybuant (F_n) standaryzowanych ZL U_n:

$$U_n = \frac{S_n - np}{\sqrt{npq}}$$

jest przy $n\to\infty$ zbieżny na całej prostej do dystrybuanty Φ standaryzowanego rozkładu normalnego N(0;1):

$$(2.10) \qquad P(U_n < u) \equiv F_n(u) \to \Phi(u), \quad u \in R.$$

Prawdziwe są również przybliżenia postaci (2.8).

PRZYKŁAD 2.3. Pr-stwo, że klient domu handlowego poszukuje ubrania o danym rozmiarze wynosi 0,2. Znajdziemy pr-stwo tego, że wśród 1500 kupujących liczba S_{1500} tych, dla których właściwy jest dany rozmiar, zawierać się będzie w przedziale (280; 330).

Wielkość S_{1500} jest ZL o rozkładzie dwumianowym b(1500;0,2). Poszukiwane pr-stwo znajdujemy na podstawie twierdzenia Moivre'a - Laplace'a i wynikającego z niego przybliżenia postaci (2.8):

$$P(280 < S_{1500} < 330) = P\left(\frac{280 - 1500\cdot 0,2}{\sqrt{1500\cdot 0,2\cdot 0,8}} < \frac{S_{1500} - 1500\cdot 0,2}{\sqrt{1500\cdot 0,2\cdot 0,8}} < \frac{330 - 1500\cdot 0,2}{\sqrt{1500\cdot 0,2\cdot 0,8}}\right) =$$

$$= P(-1{,}29 < U_{1500} < 1{,}94) \approx \Phi(1{,}94) - \Phi(-1{,}29) = 0{,}875. \quad \blacksquare$$

ZADANIA DO ROZWIĄZANIA.

1. Sprawdzić, że dla dużych $n \in N$ i dowolnego $\varepsilon > 0$ $P(|\overline{X}_n - \mu| < \varepsilon) \approx$ $\approx 2\Phi(\varepsilon\sqrt{n}/\sigma) - 1$.

2. Niech S_n oznacza ZL o rozkładzie dwumianowym b(n,p). Sprawdzić, że dla dużych n i dowolnego $\varepsilon > 0$

$$P(|\frac{S_n}{n} - p| < \varepsilon) \approx 2\Phi(\varepsilon\sqrt{n/pq}) - 1.$$

3. Przeprowadzamy $n = 300$ doświadczeń według schematu Bernoulliego z pr-stwem $p = 0{,}6$ sukcesu w pojedynczym doświadczeniu. Obliczyć pr-stwo tego, że częstość względna sukcesu S_n/n odchyli się co do modułu od pr-stwa p o nie więcej niż $\varepsilon = 0{,}05$.

4. Niech p oznacza nieznane pr-stwo, że obiekt zostanie wykryty w ciągu jednego obrotu anteny radarowej. Oszacować liczbę obrotów, przy której pr-stwo tego, że - częstość względna wykrycia obiektu będzie różnić się co do modułu od pr-stwa p o mniej niż $\varepsilon = 0{,}05$ - jest równe 0,9.

5. Wadliwość partii towaru wynosi $p = 0{,}6$. Z partii wylosowano ze zwrotem 200 sztuk towaru. Obliczyć pr-stwo, że wadliwość wśród tych 200 sztuk będzie różnić się co do modułu od wadliwości w całej partii o mniej niż $\varepsilon = 0{,}05$.

6. Pearson na 24000 "rzutów monetą" otrzymał 12012 orłów. Obliczyć pr-stwo, że przy powtórzeniu doświadczenia otrzymamy mniejsze od zaobserwowanego przez Pearsona odchylenie częstości względnej orła od pr-stwa wyrzucenia orła.

7. Pr-stwo, że dowolna ustalana linia spośród n linii w centrali telefonicznej jest zajęta, wynosi 0,6. Jak liczba n linii wystarczy, aby pr-stwo tego, że co najmniej 35% linii jest wolnych było równe 0,9.

8. Dysponujemy 80 żarówkami. Wkręcamy jedną z nich do obwodu, gdy przepali się wkręcamy następną. Zakładamy, że czas świecenia każdej z tych żarówek jest wartością ZL o rozkładzie wykładniczym z wartością oczekiwaną $\mu = 1500$h. Wyznaczyć pr-stwo, że posiadany zapas żarówek wystarczy na 100 tysięcy godzin.

9. ZL $\overline{X}$ jest średnią arytmetyczną 100 niezależnych ZL o tym samym rozkładzie pr-stwa z wartością oczekiwaną 5 i odchyleniem standardowym $\sigma = 0{,}5$. Obliczyć pr-stwo, że $\overline{X}$ przyjmie wartość z przedziału (4,9;5,1).

Odpowiedzi.

3. 0,92; 4. $n = 271$, wskazówka: $p(1-p) \le 1/4$; 5. 0,85; 6. 0,13; 7. $n = 158$.
8. 0,93; 9. 0,95.

3. PRÓBA. STATYSTYKI. PRÓBKA I JEJ OPIS.

WNIOSKOWANIE STATYSTYCZNE. W zagadnieniach związanych z rzeczywistością, w których interesujemy się pewną cechą X - zmienna losowa (ZL) i jej rozkład pr-stwa pełnią rolę modelu dla tej cechy. W większości takich praktycznych zagadnień rozkład pr-stwa ZL X i wartości jego parametrów nie są znane. Konkretne zagadnienia praktyczne dają jednak możliwość zaobserwowania wartości $x_1, x_2, \ldots, x_n$ tej cechy X. Te zaobserwowane wartości dostarczają pewnej informacji o postaci rozkładu pr-stwa tej cechy lub o wartościach jego parametrów.

Wyciąganie wniosków dotyczących nieznanej p o s t a c i lub nieznanych w a r t o ś c i p a r a m e t r ó w rozkładu pr-stwa ZL X, będącej modelem dla rozważanej cechy, na podstawie zaobserwowanych jej wartości $x_1, x_2, \ldots, x_n$, z wykorzystaniem rachunku pr-stwa, nazywa się **wnioskowaniem statystycznym**. To wnioskowanie statystyczne jest przedmiotem **statystyki matematycznej**.

PRÓBA LOSOWA. STATYSTYKI. Rozważamy pewną cechę X elementów pewnego zbioru tych elementów. Ten wyjściowy zbiór nazywa się **populacją generalną** (krócej: **populacją**). Zakładamy, że cecha X daje się traktować jako ZL.

Dokonujemy n-krotnej obserwacji ZL X. Zakładamy, że wyniki uzyskane w dowolnej obserwacji nie mają wpływu na wyniki uzyskane w pozostałych obserwacjach. Niech X_1 oznacza ZL, której wartościami są wszystkie możliwe wyniki pierwszej obserwacji, X_2 - drugiej obserwacji, ..., X_n - n-tej obserwacji. Tak określone ZL $X_1,X_2,\ldots,X_n$ są więc niezależne i mają ten sam rozkład pr-stwa co ZL X. Przyjmujemy następującą umowę.

Ciąg $(X_1,X_2,\ldots,X_n)$ ZL nazywa się **próbą losową prostą (PLP)** z populacji, w której rozważana jest cecha X (krócej: **PLP cechy X**), gdy: 1) ZL $X_1,X_2,\ldots,X_n$ są niezależne oraz 2) każda ZL X_i ma ten sam rozkład pr-stwa jak ZL X. Liczbę n nazywa się **licznością próby**. Interpretacja PLP wynika z rozważań poprzedzających jej definicję; można powiedzieć, że PLP $(X_1,X_2,\ldots,X_n)$ cechy X stanowi n-krotną "reprezentację" cechy X.

Wnioskowanie statystyczne przeprowadza się nie bezpośrednio na podstawie PLP, lecz buduje się w tym celu pewne funkcje próby. Każdą funkcję $T_n=T_n(X_1,X_2,\ldots,X_n)$ PLP $(X_1,X_2,\ldots,X_n)$ nazywa się **statystyką**. Przykłady statystyki podamy w następnych paragrafach.

Statystyka, jako funkcja ZL, jest ZL. Można wobec tego mówić o jej rozkładzie pr-stwa. Rozkład pr-stwa statystyki T_n zależy zarówno od jej postaci jak i od rozkładu ZL X, którą "reprezentują" ZL $X_1,X_2,\ldots,X_n$. Rozkład pr-stwa statystyki T_n dla u s t a l o n e g o n nazywa się **dokładnym rozkładem pr-stwa tej statystyki**. Oprócz tego rodzaju rozkładów rozważa i wykorzystuje się **graniczne rozkłady statystyk** T_n przy $n\to\infty$. Wyznaczanie rozkładów statystyk jest przeważnie trudnym zagadnieniem.

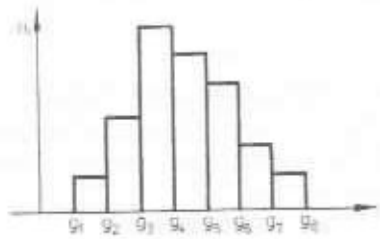
PRÓBKA I JEJ OPIS. 1. Niech $(X_1,X_2,\ldots,X_n)$ będzie PLP cechy X. Załóżmy, że w rezultacie n-krotnej obserwacji cechy X zanotowaliśmy, że: $X_1=x_1, X_2=x_2,\ldots,X_n=x_n$. Wówczas ciąg l i c z b o w y $(x_1,x_2,\ldots,x_n)$ nazywa się **realizacją** PLP $(X_1,X_2,\ldots,X_n)$ lub **zaobserwowaną próbą losową** lub po prostu **próbką** cechy X. Będziemy używać tego ostatniego terminu. Zatem: próbka $(x_1,x_2,\ldots,x_n)$ - to realizacją (wartość) PLP $(X_1,X_2,\ldots,X_n)$. Zbiór próbek nazywa się **przestrzenią prób**. Przestrzeń prób jest podzbiorem przestrzeni R^n.

2. Próbka stanowi materiał, na którym, posługując się metodami rachunku pr-stwa, przeprowadza się wnioskowanie statystyczne. Zwykle już sama próbka wymaga pewnego wstępnego opracowania. Szczegółowo zajmuje się tym **statystyka opisowa**.

W sytuacji gdy próbka $(x_1,x_2,\ldots,x_n)$ jest bardzo liczna i tym samym mało czytelna, opracowanie takie może, między innymi, polegać na zbudowaniu **szeregu rozdzielczego**, tj na utworzeniu przedziałów $<g_{i-1},g_i)$, $i=1,2,\ldots,k$, zwanych **klasami** i zaliczenie do i-tej klasy tych elementów próbki, które spełniają nierówność $g_{i-1}\le x_i<g_i$. Przyjmuje się, że *elementy tej samej klasy są równe środkowi* $\overline{x}_i=(g_{i-1}+g_i)/2$ *tej klasy*. Liczbę n_i elementów próbki, które znalazły się w i-tej klasie nazywa się **licznością** tej klasy. Zrozumiałe, że jeśli spełnione są warunki: $g_0\le\min(x_1,x_2,\ldots,x_n)$ oraz $g_k\ge\max(x_1,x_2,\ldots,x_n)$, to każdy element próbki trafia do jednej klasy i tylko jednej (zakładamy, że $g_0<g_1<g_2<\cdots<g_k$). W konsekwencji $n_1+n_2+\cdots+n_k=n$. Szereg rozdzielczy ma postać dwuwierszowej tabelki:

Klasy $g_{i-1}\div g_i$	$g_0\div g_1$	$g_1\div g_2$	...	$g_{k-1}\div g_k$
Liczności n_i	n_1	n_2	...	n_k

Obrazem geometrycznym szeregu rozdzielczego jest jego **histogram** (rys 3.1). Tak nazywa się układ k prostokątów, których podstawami są klasy $<g_{i-1},g_i>$, a wysokościami - ich liczności n_i.



Rys 3.1.

3. Do wstępnego opracowania próbki zalicza się obliczanie **charakterystyk liczbowych próbki**. Do najważniejszych charakterystyk próbki

należą: średnia arytmetyczna próbki, wariancja próbki i odchylenie standardowe próbki.

Średnią arytmetyczną próbki $(x_1,x_2,...,x_n)$ nazywamy liczbę $\overline{x}$ określoną wzorem

$$\overline{x} \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{3.1}$$

Gdy w próbce jest tylko k różnych elementów, $k<n$, z których pierwszy powtarza się n_1 razy, drugi - n_2 razy, ..., k-ty - n_k razy, to próbka taka daje się opisać tabelą

x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k

i średnią arytmetyczną takiej próbki można zapisać w postaci:

$$\overline{x} \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{k} n_i x_i, \qquad n_1+n_2+...+n_k=n. \tag{3.2}$$

Liczba $\overline{x}$ określona tym wzorem nazywa się **średnią arytmetyczną ważoną** próbki, przy tym liczba n_i nazywa się **wagą** elementu x_i. Zrozumiałe, że liczby określone wzorami (3.1) i (3.2) są równe.

Średnia arytmetyczna próbki w postaci szeregu rozdzielczego przyjmuje postać średniej arytmetycznej ważonej:

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{u} n_i\overline{x}_i, \quad \text{gdzie } \overline{x}_i=\frac{1}{2}(g_{i-1}+g_i) \text{ - środek i-tej klasy.} \tag{3.3}$$

Określenie (3.1) można zapisać w postaci $\overline{x}=\sum x_i p_i$, gdzie $p_i=$ $=\text{const}=1/n$. Widać stąd, że średnią arytmetyczną próbki można traktować jako wartość oczekiwaną ZLS o rozkładzie równomiernym. Zatem: *średnia arytmetyczna próbki ma wszystkie własności wartości oczekiwanej ZL.* W szczególności:

$$\sum_{i=1}^{n}(x_i-\overline{x})=0, \quad \overline{x}=c+\frac{1}{n}\sum_{i=1}^{n}(x_i-c), \quad \overline{x}=a\cdot\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{a}. \tag{3.4}$$

4. Za miarę rozrzutu próbki $(x_1,x_2,...,x_n)$ względem średniej arytmetycznej $\overline{x}$ najczęściej przyjmuje się **wariancję próbki** s^2 opkreśloną wzorem:

$$s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2 \tag{3.5}$$

lub **odchylenie standardowe** próbki: $s=\sqrt{s^2}$. *Wariancja próbki ma wszystkie własności wariancji ZL.* W szczególności:

$$s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-c)^2-(c-\overline{x})^2, \quad s^2=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\overline{x}^2. \tag{3.6}$$

Liczbę c dobiera się tak, aby różnice $x_i-\overline{x}$ były "wygodne" w dalszych obliczeniach. Drugi wzór (3.6) wynika z pierwszego dla $c=0$.

ZADANIA DO ROZWIĄZANIA.

1. Z jakiej własności wariancji ZL wynika pierwszy wzór (3.6). Sprawdzić również jego prawdziwość bezpośrednim rachunkiem.
2. Ze wzoru (3.6) wyznaczyć sumę $\sum(x_i-c)^2$. Dla jakiej wartości c suma ta osiąga minimum?
3. Napisać wzory (3.5) i (3.6) dla wariancji próbki: 1) w której elementy powtarzają się oraz 2) w postaci szeregu rozdzielczego.
4. Zakład produkujący wyroby elektryczne otrzymał od kooperanta 140 elementów. Wyniki pomiaru oporności (w omach) u każdego z tych elementów zgrupowano w następujący szereg rozdzielczy:

54–56	56–58	58–60	60–62	62–64	64–66	66–68	68–70	70–72
2	8	17	23	39	20	16	9	6

Wyznaczyć: a) histogram, b) średnią arytmetyczną próbki, c) wariancję próbki.

Odpowiedzi. 2. $c=\overline{x}$.

3. 1) $s^2=\frac{1}{n}\sum_{i=1}^{k}n_i(x_i-\overline{x})^2$, $s^2=\frac{1}{n}\sum_{i=1}^{k}n_i(x_i-c)^2-(c-\overline{x})^2$, $s^2=\frac{1}{n}\sum_{i=1}^{k}n_i x_i^2-\overline{x}^2$;
2) we wzorach z punktu 1) liczby x_i zastąpić liczbami $\overline{x}_i$ (środkami klas).

4. b) $\overline{x}=63{,}2$; c) $s^2=12{,}99$.

4. ESTYMATORY PUNKTOWE.

POJĘCIE ESTYMATORA. Zakładamy, że cecha X ma rozkład pr-stwa o GP f lub funkcji pr-stwa p(·) lub dystrybuancie F, które zależą od parametru θ. P o s t a ć tych funkcji jest nam znana, nie znamy natomiast w a r t o ś c i p a r a m e t r u θ; wiadomo nam tylko, że jest ona elementem danego zbioru Θ zwanego **przestrzenią parametru**. Przy tych założeniach poszukiwanie rozkładu pr-stwa cechy X sprowadza się do szacowania wartości parametru θ.

Estymacją nazywa się formę wnioskowania statystycznego polegającą na szacowaniu nieznanej wartości parametru θ rozkładu pr-stwa cechy X na podstawie próbki tej cechy. Rozróżnia się **estymację punktową** i **estymację przedziałową**. Rezultatem pierwszej jest liczba, czyli punkt na osi liczbowej; wynikiem drugiej - przedział liczbowy.

Estymatorem parametru θ nazywa się statystykę $\hat{\theta}_n$

$$\hat{\theta}_n \overset{\text{def}}{=} h(X_1, X_2, \ldots, X_n)$$

o wartościach z przestrzeni parametru Θ, której wartość $h(x_1,x_2,\ldots,x_n)$ dla zaobserwowanej próbki $(x_1,x_2,\ldots,x_n)$ przyjmuje się za wartość parametru θ:

$$h(x_1,x_2,\ldots,x_n)\approx\theta. \tag{4.1}$$

Liczbę $h(x_1,x_2,\ldots,x_n)$ w tej przybliżonej równości nazywa się **oszacowaniem** lub **oceną** parametru θ.

Podobnie określa się **estymator** (danej) **funkcji** g **parametru** θ - jest to statystyka, której wartość dla zaobserwowanej próbki przyjmuje się za wartość $g(\theta)$ funkcji g w punkcie θ.

JAKOŚĆ ESTYMATORA. Definicja estymatora dopuszcza istnienie wiele estymatorów tego samego parametru θ. Jest oczywiste, że różne estymatory parametru θ, nawet dla tej samej próbki $(x_1,x_2,\ldots,x_n)$ dostarczają różnych jego oszacowań. Powstaje wobec tego problem sensownego wyboru jednego lub kilku z nich. Racjonalność takiego wyboru osiąga się przez narzucenie pewnych warunków, jakie winien spełniać dobry estymator. Do warunków tych zaliczamy: nieobciążoność, zgodność oraz efektywność. Omówimy przede wszystkim nieobciążoność.

1.Wartość przeciętna estymatora winna być równa wartości estymowanego parametru. Ten postulat prowadzi do następnej definicji.

Estymator $\hat{\theta}_n$ parametru θ, którego wartość oczekiwana jest równa wartości estymowanego parametru,

$$E(\hat{\theta}_n)=\theta, \quad n=1,2,\ldots \tag{4.2}$$

nazywa się **estymatorem nieobciążonym** parametru θ lub **estymatorem bez błędu systematycznego**. Estymator dla którego $E(\hat{\theta}_n)\neq\theta$ nazywa się **estymatorem obciążonym** parametru θ. Przy tym różnica

$$E(\hat{\theta}_n)-\theta \tag{4.3}$$

nosi nazwę **obciążenia** lub **błędu systematycznego** estymatora.

Różnica (4.3) zwykle zależy od liczności próby n. Estymator $\hat{\theta}_n$ spełniający warunek

$$\lim_{n\to\infty}[E(\hat{\theta}_n)-\theta]=0, \quad \text{czyli} \quad \lim_{n\to\infty}E(\hat{\theta}_n)=\theta \tag{4.4}$$

nazywa się **asymptotycznie nieobciążonym estymatorem** parametru θ.

2.Estymator może być nieobciążony, a mimo to jego wartości mogą być bardzo odległe od wartości estymowanego parametru. Zrozumiałe, że z dwu nieobciążonych estymatorów lepszy jest ten, który ma mniejszy rozrzut swych wartości wokół θ. Najważniejszą miarą takiego rozrzutu jest wariancja estymatora. Zatem: *z dwu nieobciążonych estymatorów tego samego parametru lepszy jest ten, którego wariancja jest mniejsza*. Rozważania te prowadzą do następującej definicji.

Mówimy, że nieobciążony estymator $\tilde{\theta}_n$ parametru θ jest **efektywniejszy** od nieobciążonego estymatora $\hat{\theta}_n$ tego parametru, gdy $\mathrm{Var}\tilde{\theta}_n < \mathrm{Var}\hat{\theta}_n$. Powiemy, że nieobciążony estymator $\tilde{\theta}_n$ parametru θ jest **najefektywniejszy** w danym zbiorze nieobciążonych estymatorów tego parametru, gdy jest efektywniejszy od każdego estymatora z tego zbioru.

P R Z Y K Ł A D 4.1. Niech (X_1,X_2,X_3) będzie PLP cechy X. Estymujemy nieznaną wartość oczekiwaną μ tej cechy, $\mu=EX$. Spośród estymatorów:

$$\hat{\mu}_1 = X_1, \quad \hat{\mu}_2 = \frac{1}{3}(2X_1 + X_2), \quad \hat{\mu}_3 = \frac{1}{3}(2X_1 + X_2 - X_3),$$

$$\hat{\mu}_4 = \frac{1}{2}(X_1 + X_2), \quad \hat{\mu}_5 = \frac{1}{3}(X_1 + X_2 + X_3)$$

wybierzemy: a) estymatory nieobciążone parametru μ, b) najefektywniejszy estymator spośród nieobciążonych.

a) $E(\hat{\mu}_1) = \mu$, $E(\hat{\mu}_2) = \frac{1}{3}(2EX_1 + EX_2) = \frac{1}{3}(2\mu + \mu) = \mu$,

$$E(\hat{\mu}_3) = \frac{2}{3}\mu \neq \mu, \quad E(\hat{\mu}_4) = \mu, \quad E(\hat{\mu}_5) = \mu.$$

Zatem tylko $\hat{\mu}_3$ jest estymatorem obciążonym.

b) $\mathrm{Var}(\hat{\mu}_1) = \mathrm{Var}X_1 = \mathrm{Var}X = \sigma^2$,

$$\mathrm{Var}(\hat{\mu}_2) = \frac{1}{9}(4\mathrm{Var}X_1 + \mathrm{Var}X_2) = \frac{5}{9}\sigma^2,$$

$$\mathrm{Var}(\hat{\mu}_4) = \frac{1}{4}(\sigma^2 + \sigma^2) = \frac{1}{2}\sigma^2, \quad \mathrm{Var}(\hat{\mu}_5) = \mathrm{Var}\overline{X}_3 = \frac{1}{3}\sigma^2.$$

W danym zbiorze estymatorów najefektywniejszym jest więc estymator $\hat{\mu}_5 = \overline{X}_3$, czyli średnia arytmetyczna próby (X_1, X_2, X_3). ■

3. Od estymatora wymaga się aby oszacowanie (4.1) było tym lepsze, im liczność próby jest większa. W znacznym stopniu warunek ten spełnia estymator zgodny. Mówimy, że estymator $\hat{\theta}_n$ parametru θ jest **zgodny**, gdy dla każdego $\varepsilon > 0$

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| < \varepsilon) = 1. \tag{4.5}$$

Zatem estymator zgodny dla dużych liczności próby, tj dla dużych n, przyjmuje z pr-stwem bliskim jedności wartości bliskie estymowanemu parametrowi.

PRZYKŁAD 4.2. Ze słabego prawa wielkich liczb Bernoulliego (por. tw. 2.2) wynika, że częstość względem S_n/n zdarzenia A w serii n doświadczeń Bernoulliego jest zgodnym estymatorem parametru p, tj pr-stwa sukcesu A w pojedynczym doświadczeniu, $p = P(A)$. ■

ESTYMATORY WARTOŚCI OCZEKIWANEJ. Najczęściej stosowanym estymatorem $\hat{\mu}$ nieznanej wartości oczekiwanej $\mu = EX$ cechy X jest średnia arytmetyczna $\overline{X}_n$ PLP $(X_1, X_2, \ldots, X_n)$, $\hat{\mu} = \overline{X}_n$, określona wzorem:

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

O jej zaletach orzeka następujące

TWIERDZENIE 4.1. 1) Średnia arytmetyczna $\overline{X}_n$ PLP $(X_1, X_2, \ldots, X_n)$ cechy X o dowolnym rozkładzie jest nieobciążonym i zgodnym estymatorem nieznanej wartości oczekiwanej μ (jeśli ta wartość oczekiwana istnieje):

$$E\overline{X}_n = \mu \text{ oraz } \lim_{n\to\infty} P(|\overline{X}_n - \mu| < \varepsilon) = 1 \text{ przy dowolnym } \varepsilon > 0. \tag{4.5}$$

2) Gdy cecha X ma rozkład normalny $N(\mu, \sigma)$, to średnia arytmetyczna jest również najefektywniejszym estymatorem wartości oczekiwanej μ.

Dowód. Równości (4.5) wynikają bezpośrednio z twierdzeń 1.1 i 2.1, jeśli uwzględni się definicję PLP. Dowód tezy 2) pomijamy. □

Innym, stosownym estymatorem wartości oczekiwanej jest **średnia arytmetyczna wartości skrajnych**

$$\overline{X}_n = \frac{1}{2}(X_{\min} + X_{\max}), \tag{4.6}$$

gdzie $X_{\min} = \min(X_1, X_2, \ldots, X_n)$, $X_{\max} = \max(X_1, X_2, \ldots, X_n)$.

ESTYMATORY WARIANCJI. Najczęściej stosowanymi estymatorami nieznanej wariancji $\sigma^2 = \mathrm{Var}X$ cechy X są:

1) **wariancja** S^2 **PLP** $(X_1, X_2, \ldots, X_n)$ określona wzorem

$$S^2 \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2. \tag{4.7}$$

2) **skorygowana wariancja** S_1^2 **PLP** $(X_1, X_2, \ldots, X_n)$ określona wzorem:

$$S_1^2 \overset{\text{def}}{=} \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2. \tag{4.8}$$

3) **wariancja** S_0^2 **PLP** $(X_1, X_2, \ldots, X_n)$ określona wzorem

(4.9) $$S_0^2 \overset{def}{=} \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2,$$

gdy wartość oczekiwana $\mu=EX$ jest znana.

Wykazuje się, że 1) wszystkie *wariancje* S_0^2, S_1^2, S^2 *są zgodnymi estymatorami wariancji* $\sigma^2=VarX$, 2) *wariancje* S_0^2 i S_1^2 *są nieobciążonymi estymatorami wariancji* σ^2: $ES_0^2=\sigma^2$, $ES_1^2=\sigma^2$, *natomiast wariancja* S^2 *jest asymptotycznie nieobciążonym estymatorem wariacji* σ^2:

$$ES^2=\frac{n-1}{n}\sigma^2\to\sigma^2, \quad \text{gdy} \quad n\to\infty.$$

Estymatory odchylenia standardowego $\sigma=DX$, pochodne od estymatorów wariancji (4.7), (4.8) i (4.9), to: **odchylenie standardowe S PLP, skorygowane odchylenie standardowe** S_1 **PLP** oraz **odchylenie standardowe** S_0 **PLP** (gdy μ jest znane) określone wzorami:

$$S \overset{def}{=} \sqrt{S^2}, \quad S_1 \overset{def}{=} \sqrt{S_1^2}, \quad S_0 \overset{def}{=} \sqrt{S_0^2},$$

gdzie S^2, S_1^2, S_0^2 dane wzorami (4.7), (4.8), (4.9).

ZADANIA DO ROZWIĄZANIA.

1. Niech (X_1,X_2,X_3,X_4) będzie PLP ze względu na cechę X. Niech $\mu=EX$ i $\sigma^2=VarX$. Spośród statystyk

$$\hat{\mu}_1=\frac{1}{2}(X_1+X_4), \qquad \hat{\mu}_2=\frac{1}{4}(2X_1+2X_2+X_3-X_4),$$

$$\hat{\mu}_3=\frac{1}{2}(X_1+X_2+X_4), \quad \hat{\mu}_4=\frac{1}{4}(X_1+X_2+X_3+X_4)$$

wybrać: a) estymatory nieobciążone parametru μ,
b) najefektywniejszy estymator spośród nieobciążonych.

2. Znaleźć związek między statystykami S^2 i S_1^2 (por. (4.7) i (4.8)).

3. a) Wykazać nieobciążoność estymatorów wariancji: S_0^2 i S_1^2;
b) znaleźć obciążenie estymatora S^2.

4. a) Jak nazwać liczby $\bar{x}, s^2, s$ otrzymane w zadaniu 4 w paragrafie 3;
b) wyznaczyć s_1^2 i s_1 dla danych w zadaniu 4 w paragrafie 3.

5. Rozważamy masy 50-metrowych odcinków przędzy produkowanej w ustalonych, niezmiennych warunkach. Wszystkie takie odcinki stanowią populację generalną. Niech μ oznacza średnią masę takiego odcinka, czyli $\mu=EX$, gdzie X oznacza rozważaną tu cechę - masę poszczególnych odcinków. Zważono 7 takich 50-metrowych odcinków i otrzymano w (gramach): 2,73; 3,11; 2,95; 2,78; 2,85; 3,03; 2,90. Wyniki te stanowią próbkę. Na podstawie tej próbki wyznaczyć oszacowanie nieznanej wartości oczekiwanej μ, wykorzystując średnią arytmetyczną i estymator (4.6).

Odpowiedzi.

1. a) $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_4$; b) $\hat{\mu}_4$. 2. $S_1^2=\frac{n}{n-1}S^2$. 3. $ES^2-\sigma^2=\frac{-\sigma^2}{n}$.
4. $s^2\approx 12{,}99$; $s_1^2\approx 13{,}08$; $s_1\approx 3{,}62$.
5. $\mu\approx\bar{x}=2{,}91$, $\mu\approx\frac{1}{2}(x_{min}+x_{max})=2{,}92$

5. ESTYMACJA PRZEDZIAŁOWA.

POJĘCIE PRZEDZIAŁU UFNOŚCI. Rozważamy, jak w ostatnim paragrafie, cechę X o rozkładzie zależnym od parametru θ. Chcemy go estymować. Jednak tym razem z parametrem θ będziemy wiązać nie jedną statystykę $\hat{\theta}_n$ lecz parę statystyk G_1 i G_2.

Niech zatem $(X_1,X_2,\ldots,X_n)$ będzie PLP cechy X. Niech α oznacza ustaloną liczbę z przedziału (0,1), w praktyce raczej bliską zera. Gdy istnieją statystyki $G_1=h_1(X_1,X_2,\ldots,X_n;\alpha)$, $G_2=h_2(X_1,X_2,\ldots,X_n;\alpha)$ zależne również od α, takie, że

(5.1) $$P(G_1<\theta<G_2)=1-\alpha,$$

to przedział losowy (G_1,G_2) nazywa się **przedziałem ufności ze współczynnikiem ufności** $1-\alpha$ **dla parametru** θ.

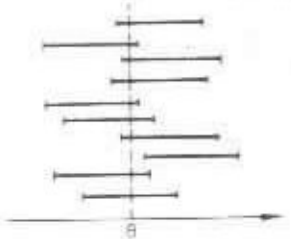
Równość (5.1) czytamy: pr-stwo pokrycia przedziałem (G_1,G_2) wartości θ jest równe $1-\alpha$. Podkreślamy: pr-stwo $1-\alpha$ w równości (5.1) należy wiązać nie z parametrem θ (który nie jest przecież ZL, lecz po prostu nieznaną stałą), ale z losowymi końcami G_1 i G_2.

Gdy dysponujemy próbką $(x_1,x_2,...,x_n)$ cechy X, to statystyki G_1 i G_2 w (5.1) przyjmują wartości

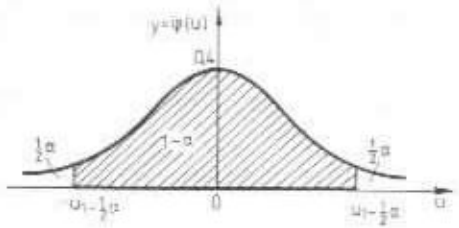
$$g_1=h_1(x_1,x_2,...,x_n;\alpha), \qquad g_2=h_2(x_1,x_2,...,x_n;\alpha);$$

wówczas przedział l i c z b o w y (g_1,g_2) nazywa się **realizacją przedziału ufności (5.1)** albo **zaobserwowanym przedziałem ufności (5.1)**.

Współczynnik ufności dotyczy bezpośrednio przedziału ufności (G_1,G_2). Nie można natomiast pisać $P(g_1<\theta<g_2)=1-\alpha$. Powstaje wobec tego pytanie o interpretację współczynnika ufności $1-\alpha$ w odniesieniu do realizacji (g_1,g_2). Odpowiedzi należy szukać w c z ę s t o ś c i o w e j i n t e r p r e t a c j i p r - s t w a z d a r z e n i a. Otóż, gdybyśmy pobrali 100 n-elementowych próbek i dla każdej z nich wyznaczyli realizację (g_1,g_2) przedziału ufności (G_1,G_2) ze współczynnikiem ufności np. $1-\alpha=0{,}95$ oraz każdorazowo twierdzili, że realizacja pokrywa parametr θ, to należy oczekiwać, że 95 razy będziemy mieć rację. Na rysunku 5.1 zaznaczono tylko 10 realizacji, z których jedna nie pokrywa parametru θ.



Rys 5.1.



Rys 5.2.

PRZEDZIAŁY UFNOŚCI DLA WARTOŚCI OCZEKIWANEJ. Zakładamy, że istnieją wartość oczekiwana $\mu=EX$ i wariancja $\sigma^2=VarX$ cechy X. *Przedziałów ufności* (G_1,G_2) *spełniających warunek* (5.1) *jest nieskończenie wiele*. W poniższych twierdzeniach będą to najkrótsze przedziały ufności o środku $\overline{X}$, a zatem najlepsze przedziały ufności (gdyż $E\overline{X}=\mu$).

TWIERDZENIE 5.1. Jeżeli cecha X ma rozkład n o r m a l n y $N(\mu,\sigma)$ o z n a n y m odchyleniu standardowym σ, to przedział

(5.2) $$(\overline{X}-u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\ \overline{X}+u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$$

jest przedziałem ufności ze współczynnikiem ufności $1-\alpha$ dla wartości oczekiwanej μ; $\overline{X}$ oznacza tu średnią arytmetyczną PLP $(X_1,X_2,...,X_n)$ cechy X, liczba $u_{1-\alpha/2}$ jest kwantylem rzędu $p=1-\alpha/2$ ZL U o rozkładzie N(0;1).

D o w ó d. Standardowe postępowanie przy wyznaczaniu przedziałów ufności polega na następujących czynnościach.

1) Wprowadzamy do rozważań "statystykę wyjściową", zależną od estymowanego parametru, której rozkład jednak nie zależy od tego parametru. W naszym przypadku jest to statystyka:

(a) $$U_n \overset{def}{=} \frac{\overline{X}-\mu}{\sigma}\sqrt{n}.$$

2) Wyznaczamy rozkład pr-stwa tej statystyki. Przy założeniach twierdzenia średnia arytmetyczna $\overline{X}$ ma rozkład normalny $N(\mu,\sigma/\sqrt{n})$, (por. tw. 1.3). ZL U_n jest zestandaryzowaną ZL $\overline{X}$, ma więc *rozkład normalny* N(0;1).

3) Dobieramy liczby a i b tak, aby

$$P(a<\frac{\overline{X}-\mu}{\sigma}\sqrt{n}<b)=1-\alpha.$$

Liczby te można wybrać na dowolną liczbę sposobów. Przedział będzie najkrótszy, jeśli $a=u_{\alpha/2}=-u_{1-\alpha/2}$ oraz $b=u_{1-\alpha/2}$, (por. rys 5.2):

$$P(-u_{1-\frac{\alpha}{2}}<\frac{X-\mu}{\sigma}\sqrt{n}<u_{1-\frac{\alpha}{2}})=1-\alpha.$$

4) Przekształcając tożsamościowo nierówność podwójną występującą po lewej stronie ostatniej równości, otrzymujemy:

$$P(\overline{X}-u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})=1-\alpha.$$

Zgodnie z definicją przedziału ufności, równość ta uzasadnia nazwanie przedziału (5.2) - przedziałem ufności ze współczynnikiem ufności $1-\alpha$ dla wartości oczekiwanej μ. □

P R Z Y K Ł A D 5.1. Niech X oznacza zużycie przędzy na wyprodukowanie jednego metra bieżącego tkaniny płaszczowej. Dokonano $n=9$ niezależnych, jednakowo dokładnych pomiarów i otrzymano wyniki

(w [g/mb]) x_i: 464, 468, 473, 474, 476, 480, 482, 487, 489. Zakładając, że rozważana tu cecha X ma rozkład normalny $N(\mu;8)$ i przyjmując współczynnik ufności $1-\alpha=0,95$, wyznaczymy realizację przedziału ufności dla przeciętnego zużycia przędzy na wyprodukowanie jednego metra tkaniny, tj dla $\mu=EX$.

Spełnione są tu założenia twierdzenia 5.1. Szukaną realizacją przedziału ufności (5.2) jest przedział liczbowy

$$(g_1,g_2), \quad \text{gdzie} \quad g_1=\overline{x}-d, \quad g_2=\overline{x}+d, \quad d=u_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}.$$

Obliczamy najpierw średnią arytmetyczną próbki:

$$\overline{x}=\frac{1}{9}\sum_{i=1}^{9}x_i=475+\frac{1}{9}\sum_{i=1}^{9}(x_i-475)=475+\frac{1}{9}\cdot 18=477.$$

Z tablic kwantyli rozkładu normalnego N(0;1) (tablica Ib na końcu książki) odczytujemy kwantyl $u_{1-\alpha/2}=u_{0,975}=1,96$. Zatem:

$$d=1,96\cdot 8/3\approx 5,2, \quad g_1=477-5,2=471,8, \quad g_2=477+5,2=482,2.$$

Szukaną realizacją jest więc przedział liczbowy (471,8; 482,2). ■

Sytuacje, gdy znane jest odchylenie standardowe σ cechy X, występują rzadko. Dlatego przydatne jest następujące

TWIERDZENIE 5.2. Jeżeli cecha X ma rozkład n o r m a l n y $N(\mu,\sigma)$, lecz odchylenie standardowe σ jest n i e z n a n e, to przedział

$$(5.3) \qquad \left(\overline{X}-t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}, \ \overline{X}+t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n-1}}\right)$$

jest przedziałem ufności ze współczynnikiem ufności $1-\alpha$ dla wartości oczekiwanej μ; $\overline{X}$ i S oznaczają tu średnią arytmetyczną i odchylenie standardowe PLP $(X_1,X_2,\ldots,X_n)$ cechy X, liczba $t_{n-1,1-\alpha/2}$ jest kwantylem rzędu $p=1-\alpha/2$ ZL t_{n-1} o rozkładzie Studenta z $n-1$ stopniami swobody.

W dowodzie tego twierdzenia rolę statystyki wyjściowej pełni **statystyka Studenta** t_{n-1} określona równością

$$(5.4) \qquad t_{n-1}=\frac{\overline{X}-\mu}{S}\sqrt{n-1}.$$

Dowodzi się, że statystyka ta ma rozkład Studenta z $n-1$ stopniami swobody. Dalej dowód przebiega analogicznie jak twierdzenia 5.1.

P R Z Y K Ł A D 5.2. Dane jak w przykładzie 5.1. Jednak tym razem nie zakładamy znajomości odchylenia standardowego σ. Przyjmując współczynnik ufności $1-\alpha=0,95$ wyznaczymy realizację przedziału ufności dla wartości μ przeciętnego zużycia przędzy na 1 mb tkaniny.

Spełnione są założenia ostatniego twierdzenia. Szukaną realizacją przedziału ufności (5.3) jest przedział liczbowy

$$(g_1,g_2), \quad \text{gdzie} \quad g_1=\overline{x}-d, \quad g_2=\overline{x}+d, \quad d=t_{n-1,1-\frac{\alpha}{2}}\frac{s}{\sqrt{n-1}}.$$

Liczbę $\overline{x}=477$ znaleźliśmy już w poprzednim przykładzie. Rachunki potrzebne dla obliczenia wariancji próbki należy przeprowadzać w tabeli o kolumnach: x_i, $x_i-\overline{x}$, $(x_i-\overline{x})^2$. Otrzymujemy:

$$s^2=\frac{1}{9}\sum_{i=1}^{9}(x_i-\overline{x})^2=\frac{1}{9}\cdot 554\approx 61,56; \quad s\approx 7,85.$$

Z tablic kwantyli rozkładu Studenta odczytujemy kwantyl

$$t_{n-1,1-\alpha/2}=t_{8;0,975}=2,306.$$

Zatem $d=2,306\frac{7,8}{\sqrt{8}}\approx 6,4$; $g_1=477-6,4=470,6$; $g_2=477+6,4=483,4$.

Szukaną realizacją jest więc przedział liczbowy (470,6; 483,4).

Zauważmy, że tym razem otrzymany przedział jest szerszy, a więc gorszy (środek ten sam, współczynnik ten sam). Jest to zrozumiałe: tym razem dysponowaliśmy mniejszą informacją o procesie zużycia przędzy w procesie produkcji, bo nie znaliśmy odchylenia standardowego σ. ■

Nie zawsze można przyjąć, jak do tej pory, że cecha X ma rozkład normalny. W takich przypadkach, gdy ponadto dysponujemy dużą próbką, przydatne może być następujące

TWIERDZENIE 5.3. Jeżeli cecha X ma d o w o l n y rozkład i istnieją jej wartość oczekiwana i wariancja (choć nie są znane) oraz liczność n próby jest duża ($n\geq 30$), to przedział

$$(5.5) \qquad \left(\overline{X}-u_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}, \ \overline{X}+u_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}\right)$$

jest przedziałem ufności ze współczynnikiem ufności $1-\alpha$ dla wartości oczekiwanej μ; liczba $u_{1-\alpha/2}$ jest tu kwantylem rzędu $p=1-\alpha/2$ ZL U o rozkładzie N(0;1), $\overline{X}$ oznacza średnią arytmetyczną PLP $(X_1,X_2,\ldots,X_n)$

cechy X, s - jest zaobserwowaną wartością odchylenia standardowego S tej PLP.

D o w ó d . Zgodnie z twierdzeniem Linderberga-Lévyego ZL

$$U_n=\frac{S_n-n\mu}{\sigma\sqrt{n}}\equiv\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$$

dla dużych n ma w przybliżeniu rozkład normalny N(0;1) nawet wtedy, gdy nieznane odchylenie standardowe σ zastąpimy jego oszacowaniem s z danej dużej próbki , $\sigma\approx s$:

$$U_n=\frac{\overline{X}-\mu}{s/\sqrt{n}}.$$

Przyjmując tę ZL U_n za "statystykę wyjściową" i postępując identycznie jak w dowodzie twierdzenia 5.1, otrzymujemy przedział (5.5).

P R Z Y K Ł A D 5.3. Niech X oznacza czas ([godz]) bezawaryjnej pracy wyrobu (np. całej maszyny, jej podzespołu lub pojedynczego elementu) produkowanego masowo. Interesuje nas wartość oczekiwana $\mu=EX$. Dokonano obserwacji czasu bezawaryjnej pracy n=60 sztuk wyrobu i obliczono $\overline{x}=845$, s=42. Wyznaczymy realizację przedziału ufności dla wartości oczekiwanej μ czasu bezawaryjnej pracy wyrobu, przyjmując $1-\alpha=0{,}90$.

Mamy tu: rozkład cechy X dowolny, liczność n próbki duża. Przedziałem ufności jest zatem przedział (5.5), a jego realizacją przedział liczbowy

$$(g_1,g_2),\quad \text{gdzie}\quad g_1=\overline{x}-d,\ g_2=\overline{x}+d,\quad d=u_{1-\frac{\alpha}{2}}\frac{s}{\sqrt{n}}.$$

Z tablic kwantyli rozkładu normalnego odczytujemy kwantyl $u_{1-\alpha/2}=u_{0,95}=1{,}65$. Obliczamy:

$$d=1{,}65\frac{42}{\sqrt{60}}\approx 8{,}9,\quad g_1=845-8{,}9=836{,}1\approx 836,\quad g_2=845+8{,}9\approx 854.$$

Zatem: szukaną realizacją 90-procentowego przedziału ufności dla nieznanej wartości przeciętnej czasu bezawaryjnej pracy produkowanego wyrobu jest przedział liczbowy (836; 854). ■

ZADANIA DO ROZWIĄZANIA.

1. a) Wyznaczyć długość przedziału ufności (5.2); b) czy długość ta jest ZL? c) Jak zmienia się ta długość, gdy: 1) zwiększamy współczynnik ufności, 2) zwiększamy liczność próbki; d) jaki wpływ na długość przedziału ufności ma odchylenie standardowe σ ?
2. Znaleźć środek przedziałów ufności (5.2), (5.3), (5.5). Wyznaczyć wartość oczekiwaną znalezionego środka. Czy otrzymamy wynik jest korzystny dla tych przedziałów?
3. Na podstawie danych z zadania 5 w paragrafie 4 zakładając, że rozkład masy odcinków przędzy jest normalny $N(\mu,\sigma)$ i przyjmując współczynnik ufności $1-\alpha=0{,}95$, wyznaczyć realizację przedziału ufności dla przeciętnej masy μ.
4. Treść jak w zadaniu 3 z tym, że odchylenie standardowe traktujemy jako znane z uprzednich, bardziej szczegółowych badań: $\sigma=0{,}124$.
5. W rezultacie pomiarów twardości stali, otrzymano (w umownych jednostkach): 13,1; 12,8; 11,9; 12,4; 13,5; 13,7; 12,0; 13,8; 10,6; 12,4; 13,5; 11,7; 13,9; 11,5; 12,5; 11,9. Zakładając, że pomiary te są wartościami ZL X o rozkładzie $N(\mu,\sigma)$ i przyjmując współczynnik ufności $1-\alpha=0{,}90$, wyznaczyć realizację przedziału ufności dla nieznanej przeciętnej twardości μ cechy X.
6. Treść jak w zadaniu 5 z tym, że odchylenie standardowe traktujemy jako znane z uprzednich, bardziej szczegółowych badań $\sigma=0{,}90$.
7. Dokonano 100 pomiarów grubości przędzy i zgrupowano je w sześciu klasach o tej samej szerokości; pierwszą klasą jest przedział <8,10); ostatnią - przedział (18,20>; liczności klas są odpowiednio równe: 5, 21, 35, 24, 8, 7. Przyjmując współczynnik ufności $1-\alpha=0{,}95$ wyznaczyć realizację przedziału ufności dla przeciętnej grubości μ przędzy.
8. Na podstawie danych z zadania 4 w paragrafie 3, przyjmując współczynnik ufności $1-\alpha=0{,}95$, wyznaczyć realizację przedziału ufności dla przeciętnej oporności μ.

O d p o w i e d z i .

3. (2,78; 3,04); 4. (2,81; 3,01); 5. (12,16; 12,99); 6. (12,20; 12,94);
7. (13,12; 14,08); 8. (62,6; 63,8).

6. WSTĘPNE POJĘCIA WERYFIKACJI HIPOTEZ.

W tym i dwu następnych paragrafach krótko omówimy drugą, obok estymacji, formę wnioskowania statystycznego - weryfikację hipotez.

HIPOTEZY STATYSTYCZNE. **Hipotezą statystyczną** (krócej: **hipotezą**) nazywa się każde przypuszczenie dotyczące nie w pełni znanego rozkładu pr-stwa cechy X, gdy prawdziwość tego przypuszczenia jest weryfikowana na podstawie próbki $(x_1,x_2,...,x_n)$ tej cechy X.

Hipotezę, która orzeka o wartościach parametrów rozkładu cechy X, gdy znana jest postać tego rozkładu (lub dostępna jest próbka o dużej liczności), nazywa się **hipotezą parametryczną**. Hipotezę, która nie jest parametryczna nazywamy hipotezą **nieparametryczną**.

Ze względu na to, w jakim stopniu hipoteza określa rozkład cechy X, rozróżnia się hipotezy proste i hipotezy złożone. Hipoteza, która całkowicie określa rozkład pr-stwa cechy X nazywa się **hipotezą prostą**. Hipoteza, która nie jest prosta nazywa się **hipotezą złożoną**.

PRZYKŁAD 6.1. a) Niech X oznacza liczbę zgłoszeń w centrali telefonicznej w przedziale czasu o długości t. Z dotychczasowych badań wiadomo, że cecha X ma rozkład Poissona z parametrem λt, ale wartość λ nie jest znana. Stawiamy przypuszczenie: $\lambda=3$. Przypuszczenie to jest hipotezą statystyczną $H(\lambda=3)$; jest to prosta hipoteza parametryczna.

b) Dane jak wyżej. Stawiamy przypuszczenie: $\lambda>3$. Przypuszczenie to jest złożoną hipotezą parametryczną.

c) Niech X oznacza liczbę interesantów w ciągu godziny w kasie biletowej. Z dotychczasowych, licznych obserwacji wiadomo, że w ciągu godziny zgłasza się przeciętnie 20 interesantów. Stawiamy hipotezę H (X ma rozkład Poissona). Hipoteza ta dotyczy postaci rozkładu, którego wartość przeciętna jest znana $EX=\lambda=20$. Jest to prosta hipoteza nieparametryczna.

d) Cecha X jak w punkcie c). Wiadomo, że $EX>20$. Stawiamy hipotezę H (cecha X ma rozkład Poissona). Jest to złożona hipoteza nieparametryczna. ■

Zwykle rozważa się dwie wzajemnie wykluczające się hipotezy: hipotezę zerową H_0 i hipotezę alternatywną H_a. W przypadku hipotez parametrycznych w roli hipotezy alternatywnej H_a obieramy tę z dwu hipotez, dla której pragniemy potwierdzenia prawdziwości. Natomiast hipoteza zerowa H_0 pełni rolę "hipotezy roboczej", tj hipotezy, która będzie podlegać bezpośredniej weryfikacji. Tylko jedna z dwu hipotez H_0 i H_a jest prawdziwa, dlatego odrzucenie hipotezy H_0 jest argumentem za przyjęciem hipotezy alternatywnej H_a. Najczęściej hipoteza zerowa jest hipotezą prostą.

Formułowanie hipotezy zerowej i alternatywnej ilustruje

PRZYKŁAD 6.2. Oprócz innych "pozastatystycznych" argumentów, chcemy potwierdzenia procedurami statystycznymi, że "przeciętna wytrzymałość θ materiału wyprodukowanego według nowej technologii jest większa od wytrzymałości θ_0 starego materiału". To przekonanie o wyższości nowego materiału czynimy treścią hipotezy alternatywnej $H_a(\theta>\theta_0)$. Hipotezę zerową formułujemy następująco: $H_0(\theta=\theta_0)$. ■

TEST STATYSTYCZNY. Weryfikację hipotezy przeprowadza się za pomocą testu statystycznego (krócej: testu).

Zakładamy, że dysponujemy próbką $(x_1,x_2,...,x_n)$ cechy X; hipoteza H_0 dotyczy rozkładu pr-stwa tej cechy. **Testem hipotezy** H_0 nazywamy regułę postępowania, która próbce $(x_1,x_2,...,x_n)$ cechy X przyporządkowuje jedną z dwu decyzji: decyzję d_0 - odrzucić hipotezę H_0 (jako fałszywą), decyzję d_0^* - przyjąć hipotezę H_0 (jako prawdziwą).

Test, którym weryfikuje się hipotezę parametryczną (nieparametryczną) nazywa się testem **parametrycznym** (**nieparametrycznym**). Test nieparametryczny, którym weryfikuje się hipotezą o postaci rozkładu, nazywa się **testem zgodności**.

W wyniku weryfikacji hipotezy H_0 danym testem możemy podjąć poprawną decyzję, ale też (mimo poprawnego testowania) narażeni jesteśmy na popełnienie błędów dwojakiego rodzaju:

1) **błędu pierwszego rodzaju** - odrzucenie hipotezy H_0, gdy w rzeczywistości (nam nieznanej) jest ona prawdziwa,

2) **błędu drugiego rodzaju** - przyjęcie hipotezy H_0, gdy w rzeczywistości jest ona fałszywa.

Niestety, podejmując w procesie weryfikacji decyzje d_0 albo d_0' - nie wiemy, czy popełniliśmy błąd. Dlatego nasza kontrola błędów I i II rodzaju polega nie na ich eliminacji, lecz na minimalizacji pr-stw popełnienia tego błędu. Przy dokładniejszej analizie okazuje się jednak, że gdy liczność próbki jest ustalona, to nie jest możliwa minimalizacja pr-stw popełnienia każdego z tych błędów; przy tym kontrola pr-stw popełnienia błędu II rodzaju jest trudna. Konieczny tu wobec tego kompromis polega na: 1) kontroli pr-stw popełnienia błędu I rodzaju oraz 2) rezygnacji z kontroli pr-stw popełnienia błędu II rodzaju. W konsekwencji musimy wtedy rezygnować z podejmowania decyzji d_0' - przyjmowania hipotezy zerowej H_0.

Takie "niepełne" testy, w których rezygnuje się z podejmowania decyzji przyjmowania hipotezy zerowej, nazywa się **testami istotności**. Wtedy zamiast decyzji d_0' ograniczamy się do stwierdzenia d_0: "d a n a p r ó b k a n i e p r z e c z y h i p o t e z i e H_0" albo "mocniejszego" stwierdzenia "nie ma podstaw statystycznych do odrzucenia hipotezy H_0".

Zauważmy, że w rezultacie stosowania testu istotności narażeni jesteśmy na popełnienie tylko błędu I rodzaju. Kontrola tego błędu polega na kontroli pr-stwa popełnienia tego błędu. Z góry zadaną małą liczbę $\alpha \in (0,1)$ (zwykle $\alpha=0{,}05$, $\alpha=0{,}01$) taką, że pr-stwo popełnienia błędu I rodzaju nie przekracza tej liczby

(6.1) $$P(\text{błąd I rodzaju}) \le \alpha$$

nazywa się **poziomem istotności** testu hipotezy H_0; litera P w tej nierówności oznacza rozkład pr-stwa zadany hipotezą H_0).

Test istotności hipotezy H_0 przeciwko hipotezie alternatywnej H_a na poziomie istotności α zwykle przebiega według następujących trzech etapów.

1) O b i e r a m y tzw. **statystykę testową** T_n i o b l i c z a m y jej wartość $t_{n,emp}$ dla danej próbki $(x_1, x_2, \ldots, x_n)$.

2) Przy założeniu prawdziwości hipotezy zerowej H_0, w y z n a c z a m y r o z k ł a d p r - s t w a s t a t y s t y k i T_n.

3) P o s t ę p o w a n i e d e c y z y j n e. Z tablic rozkładu statystyki T_n odczytujemy kwantyl $t_{n,p}$, gdzie p jest jedną z liczb:

α, $\alpha/2$, $1-\alpha$, $1-\alpha/2$. Za pomocą tych kwantyli (przy weryfikacji hipotez nazywa się je **wartościami krytycznymi**) budujemy tzw. **zbiór krytyczny** K_α, mający następującą własność:

a) jeśli $t_{n,emp} \in K_\alpha$, to na poziomie istotności α podejmujemy decyzję d_0: odrzucamy hipotezę H_0.

b) gdy $t_{n,emp} \notin K_\alpha$, to ograniczymy się do stwierdzenia d_0: dana próbka nie przeczy hipotezie H_0 na danym poziomie istotności α.

Omówimy nieco dokładniej to postępowanie.

Ad 1. Statystyka testowa T_n winna spełniać m.in. następujące wymagania: a) musi stanowić porównanie charakterystyki cechy X i odpowiedniej charakterystyki PLP; może to być np. różnica $\overline{X}-\mu$, gdy hipoteza H_0 dotyczy wartości oczekiwanej $\mu = EX$; b) winna uwzględnić informacje dotyczące cechy X znane jeszcze przed pobraniem próbki; c) musi być możliwe, przy założeniu prawdziwości hipotezy H_0, wyznaczenie jej rozkładu pr-stwa dokładnego lub granicznego.

Ad 2. W tym wykładzie nie będziemy (z jednym wyjątkiem) wyznaczać rozkładu statystyki testowej T_n - będziemy powoływać się na jej rozkład; będą to trzy rozkłady omówione w paragrafie 1.

Ad 3. Warunek (6.1) można zapisać w postaci:

(6.2) $$P(T_n \in K_\alpha | H_0 \text{ jest prawdziwa}) \le \alpha.$$

Zbiór krytyczny K_α testu hipotezy zerowej na danym poziomie istotności spełnia więc warunek: *pr-stwo przyjęcia przez statystykę testową wartości z tego zbioru, przy założeniu prawdziwości hipotezy zerowej, nie przekracza poziomu istotności (w przypadku prostej hipotezy zerowej i rozkładu ciągłego jest mu równe).*

Zbiór krytyczny K_α spełniający warunek (6.2) można wybrać na wiele sposobów (bo ta sama liczba α może być pr-stwem wielu różnych zdarzeń). Powstaje wobec tego problem wyboru najlepszego z nich. Powiemy, że zbiór krytyczny K_α jest **najlepszy**, jeśli dla każdego innego zbioru krytycznego K'_α

$$P(T_n \in K_\alpha | H_a \text{ jest prawdziwa}) \ge P(T_n \in K'_\alpha | H_a \text{ jest prawdziwa}),$$

czyli, gdy najczęściej prowadzi do odrzucenia hipotezy zerowej, gdy w rzeczywistości prawdziwa jest hipoteza alternatywna (a więc prowadzi do

trafnej decyzji). Rozważane dalej testy, to testy o najlepszych zbiorach krytycznych.

Uzasadnimy jeszcze racjonalność decyzji 3a, tj odrzucenie hipotezy H_0, gdy $T_{n,emp.} \in K_\alpha$. Otóż, jeśli hipoteza H_0 jest prawdziwa, to zgodnie z (6.2), pr-stwo zdarzenia $T_n \in K_n$ nie przekracza poziomu istotności α, a więc jest bardzo małe i nie powinno się zrealizować dla danej próbki $(x_1,x_2,...,x_n)$. Ponieważ jednak dla danej próbki zdarzenie $T_n \in K_\alpha$ zrealizowało się, więc jesteśmy skłonni wnioskować, że założenie, przy którym pr-stwo $P(T_n \in K_\alpha | H_0$ jest prawdziwa) zostało obliczone, a więc założenie, że hipoteza H_0 jest prawdziwa - jest nieprawdziwe. Dlatego odrzucamy hipotezę H_0. Możemy się przy tym mylić, ale pr-stwo takiej pomyłki (pr-stwo błędu I rodzaju) nie przekracza poziomu istotności, jest więc małe.

ZADANIA DO ROZWIĄZANIA.

1. W konkursie na obsadzenie kierowniczego stanowiska zgłasza się kandydat A. Jaki rodzaj błędu popełnia się: a) odrzucając niesłuszną hipotezę H_0(kandydat nadaje się na to stanowisko), b) przyjmując błędnie hipotezę H_0?
2. Niech μ oznacza wartość oczekiwaną cechy X, $\mu = EX$. Załóżmy, że osoba A weryfikuje hipotezę zerową $H_0^A(\mu \le 5)$ przeciwko hipotezie alternatywnej $H_a^A(\mu > 5)$, zaś osoba B weryfikuje hipotezę zerową $H_0^B(\mu > 5)$ przeciwko hipotezie alternatywnej $H_a^B(\mu \le 5)$.
   a) Na czym będzie polegać błąd I rodzaju (I^A) i błąd II rodzaju (II^A) dla osoby A?
   b) Na czym będzie polegać błąd I rodzaju (I^B) i błąd II rodzaju (II^B) dla osoby B?
   c) Porównać błędy I^A i II^B oraz I^B i II^A. Wyciągnąć wniosek ogólny.
3. Gdy $t_{n,emp.} \notin K_\alpha$, to jak wiemy, ograniczamy się do stwierdzenia d_0: dana próbka nie przeczy hipotezie H_0. Uzasadnić potrzebę tak ostrożnego orzekania w tym przypadku.

Odpowiedzi.

1. a) błąd I rodzaju; b) błąd II rodzaju. 2. c) błędy I^A i II^B są identyczne, błędy I^B i II^A są identyczne. To czy błąd jest I czy II rodzaju zależy od postaci hipotezy zerowej. 3. Jest praktycznie pewne (bo pr-stwo tego jest $\ge 1-\alpha$), że prawdziwa jest implikacja: "jeśli hipoteza H_0 jest prawdziwa, to zaobserwujemy relację $t_{n,emp.} \notin K_\alpha$". W przypadku zaobserwowania relacji $t_{n,emp.} \notin K_\alpha$ podjęcie decyzji przyjęcia hipotezy H_0 (w przekonaniu, że jest prawdziwa), oznacza traktowanie implikacji odwrotnej do implikacji prawdziwej jako również implikacji prawdziwej, co jest błędem.

7. TESTY DLA WARTOŚCI OCZEKIWANEJ.

Nie znamy wartości oczekiwanej μ cechy X, $\mu = EX$. Jesteśmy zainteresowani w potwierdzeniu testami statystycznymi jednej z hipotez

$$H_a(\mu > \mu_0), \quad H_a(\mu < \mu_0), \quad H_a(\mu \ne \mu_0),$$

gdzie μ_0 jest dane. Rolę hipotezy zerowej pełni hipoteza $H_0(\mu = \mu_0)$. W całym paragrafie α oznacza poziom istotności, $(X_1,X_2,...,X_n)$ jest PLP cechy X. Omówimy 9 testów przy następujących warunkach.

Warunki I (WI): 1) cecha X ma rozkład normalny $N(\mu,\sigma)$, 2) odchylenie standardowe σ jest znane, 3) liczność n próbki $(x_1,x_2,...,x_n)$ - dowolna.

Warunki II (WII): 1) cecha X ma rozkład normalny $N(\mu,\sigma)$, 2) odchylenie standardowe σ jest nieznane, 3) liczność n próbki - dowolna.

Warunki III (WIII): 1) cecha X ma rozkład dowolny o wartości oczekiwanej μ i odchyleniu standardowym σ, 2) odchylenie standardowe σ jest nieznane, 3) liczność n próbki - duża ($n \ge 30$).

TEST HIPOTEZY $H_0(\mu = \mu_0)$ PRZECIWKO HIPOTEZIE $H_a(\mu > \mu_0)$ W WARUNKACH WI. Test ten, zgodnie z ogólnymi rozważaniami w ostatnim paragrafie, przebiega według trzech następujących etapów.

1) Określamy statystykę testową U

$$U=\frac{\overline{X}-\mu_0}{\sigma}\sqrt{n} \tag{7.1}$$

i obliczamy jej wartość u_{emp} dla danej próbki $(x_1,x_2,\ldots,x_n)$. Zauważmy, że statystyka U mierzy "rozbieżność" (w tym przypadku różnicę) między wartością $\overline{x}$ statystyki $\overline{X}$ ($\overline{X}$ jest dobrym estymatorem wartości oczekiwanej μ) i wartością μ_0 parametru μ zadaną hipotezą H_0; ponadto uwzględnia znajomość odchylenia standardowego σ.

2) Przy założeniu prawdziwości hipotezy $H_0(\mu=\mu_0)$, statystyka (7.1) ma standaryzowany rozkład normalny N(0,1). Wynika to z definicji PLP i tw. 1.3, jeśli zauważyć, że przekształcenia po prawej stronie wzoru (7.1) stanowią standaryzowanie średniej arytmetycznej $\overline{X}$.

3) Z tablic rozkładu N(0,1) odczytujemy kwantyl $u_{1-\alpha}$ i budujemy prawostronny zbiór krytyczny

$$K_\alpha=<u_{1-\alpha},\infty). \tag{7.2}$$

a) Jeśli $u_{emp}\in K_\alpha$, to na poziomie istotności α odrzucamy hipotezę zerową $H_0(\mu=\mu_0)$ na korzyść hipotezy alternatywnej $H_a(\mu>\mu_0)$. Uzasadnienie tej decyzji znajdzie Czytelnik w końcowej części paragrafu 6.

b) Gdy $u_{emp}\notin K_\alpha$, to ograniczamy się do stwierdzenia: dana próbka nie przeczy hipotezie $H_0(\mu=\mu_0)$, (uzasadnienie - por. odp. do zadania 3 w paragrafie 6).

Uzasadnimy jeszcze postać (7.2) zbioru krytycznego. Bierzemy tu prawostronny zbiór krytyczny, tj zbiór postaci $K_\alpha=<b,\infty)$, ponieważ duże wartości u_{emp} statystyki testowej U świadczą przeciwko prawdziwości hipotezy $H_0(\mu=\mu_0)$ i są jednocześnie argumentem za prawdziwością hipotezy alternatywnej $H_a(\mu>\mu_0)$, w potwierdzeniu prawdziwości której jesteśmy zainteresowani. Dlaczego $b=u_{1-\alpha}$? Otóż ze związku (6.2), w którym zakładamy znak równości, otrzymujemy:

$$P(U\in K_\alpha|H_0 \text{ jest prawdziwa})=P(U\in<b,\infty)|H_0 \text{ jest prawdziwa})=$$
$$=P(U\geq b|H_0 \text{ jest prawdziwa})=\alpha.$$

Stąd

$$P(U<b|H_0 \text{ jest prawdziwa})=1-\alpha.$$

Z założenia, że H_0 jest prawdziwa wynika, że statystyka U ma rozkład normalny N(0,1). Stąd i z ostatniej równości otrzymujemy, że b jest kwantylem rzędu $p=1-\alpha$ ZL U o rozkładzie N(0,1): $b=u_{1-\alpha}$.

PRZYKŁAD 7.1. Według normy przeciętny czas wykonania danego detalu wynosi 28 godzin. Pracownicy wyrażali obawę, że czas ten jest większy. Dokonano pomiaru czasu wykonania 15 detali otrzymując wyniki x_i: 27, 30, 27, 31, 29, 31, 27, 31, 29, 31, 29, 31, 27, 30, 28. Zakładając, że wyniki te są wartościami ZL X o rozkładzie normalnym N(μ;1,7) zbadamy, czy przypuszczenia pracowników o zaniżeniu normy dadzą się potwierdzić statystycznie.

Czas wykonania detalu według normy uważamy za hipotezę zerową $H_0(\mu=28)$, a wysuwane obawy pracowników o zaniżeniu normy czynimy treścią hipotezy alternatywnej $H_a(\mu>28)$. Obieramy poziom istotności $\alpha=0{,}02$. Warunki zadania czynią zadość warunkom W1. Stosujemy omówiony wyżej test.

1) Statystyka testowa jest postaci (7.1). Dla danej próbki obliczamy jej wartość empiryczną, czyli u_{emp}:

$$u_{emp}=\frac{\overline{x}-\mu_0}{\sigma}\sqrt{n}.$$

Ponieważ

$$\overline{x}=\frac{1}{15}\sum_{i=1}^{15}x_i=28+\frac{1}{15}\sum_{i=1}^{15}(x_i-28)=28+\frac{1}{15}18=29{,}2,$$

więc

$$u_{emp}=\frac{29{,}2-28}{1{,}7}\sqrt{15}\approx 2{,}73.$$

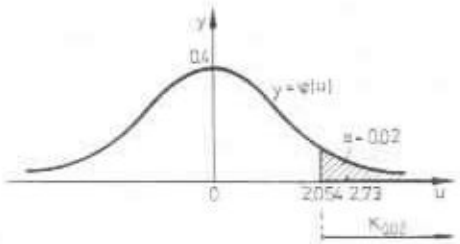
2) Statystyka U określona równością (7.1), przy założeniu prawdziwości hipotezy $H_0(\mu=28)$ ma rozkład normalny N(0,1).

3) Z tablic kwantyli tego rozkładu odczytujemy kwantyl $u_{1-\alpha}=u_{0{,}98}=2{,}05$ i budujemy prawostronny (dlaczego) zbiór krytyczny (por. rys 7.1)

$$K_{0{,}02}=<2{,}05;\infty).$$

Ponieważ $u_{emp.}=2{,}73\in K_{0{,}02}=<2{,}05;\infty)$, więc na poziomie istotności $\alpha=0{,}02$ odrzucamy hipotezę zerową $H_0(\mu=28)$ na korzyść hipotezy alternatywnej $H_a(\mu>28)$ orzekającej, że przeciętna czasochłonność wy-

konania jednego detalu jest większa od 28 godzin; zatem rozważania statystyczne wskazują (ale nie muszą być rozstrzygające), że norma wymaga korekty w kierunku sygnalizowanym przez pracowników. ■



Rys 7.1.

DALSZE TESTY DLA WARTOŚCI OCZEKIWANEJ. Zauważmy, że w omówionym teście podstawowe "elementy", to: postać hipotezy alternatywnej, postać statystyki testowej i postać zbioru krytycznego. Dla uniknięcia powtórzeń pozostałe 8 testów dla wartości oczekiwanej opisujemy w tabeli na następnej stronie. W tabeli tej opisany jest również omówiony już test.

PRZYKŁAD 7.2. Proces produkcji, którego ostatnim elementem jest pakowarka, uważa się za uregulowany, gdy średnia waga paczek wynosi 500g. Po 8 godzinach pracy proces produkcji może wymagać regulacji. Dlatego co 8 godzin pracownik na podstawie próbki losowej, podejmuje w tej sprawie decyzję. Stawiając się w roli tego pracownika zdecydujemy, czy potrzebna jest regulacja w przypadku zaobserwowania następującej próbki (w gramach): 506, 502, 498, 501, 507, 506, 496, 503, 497, 504. Przyjmujemy poziom istotności $\alpha=0{,}05$. Zakładamy poza tym, że zaobserwowane wyniki są wartościami ZI. o rozkładzie normalnym.

Populacją generalną jest tu zbiór wszystkich paczek, cechą X - waga tych paczek. Jako hipotezę zerową przyjmujemy, że proces produkcji nie rozregulował się, czyli $H_0(\mu=500)$, gdzie μ oznacza wartość przeciętną cechy X, $\mu=EX$. Ponieważ znaczne odchylenia wagi paczek od 500g, zarówno w górę jak i w dół, nie są wskazane, więc hipotezę alternatywną formułujemy w postaci $H_a(\mu\neq500)$. Warunki zadania

Nr testu	Warunki	H_a	Statystyka testowa	Rozkład stat.testowej	Zbiór krytyczny K_α
1a	W I	$\mu>\mu_0$	$U=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$	Normalny N(0,1)	$<u_{1-\alpha},\infty)$
1b		$\mu<\mu_0$			$(-\infty,-u_{1-\alpha}>$
1c		$\mu\neq\mu_0$			$(-\infty,-u_{1-\frac{\alpha}{2}}>\cup<u_{1-\frac{\alpha}{2}},\infty)$
2a	W II	$\mu>\mu_0$	$t_{n-1}=\frac{\overline{X}-\mu_0}{S/\sqrt{n-1}}$	Studenta o $k=n-1$ stop. swobody	$<t_{n-1,1-\alpha},\infty)$
2b		$\mu<\mu_0$			$(-\infty,-t_{n-1,1-\alpha}>$
2c		$\mu\neq\mu_0$			$(-\infty,-t_{n-1,1-\frac{\alpha}{2}}>\cup<t_{n-1,1-\frac{\alpha}{2}},\infty)$
3a	W III	$\mu>\mu_0$	$U=\frac{\overline{X}-\mu_0}{s/\sqrt{n}}$	W przybliżeniu normalny N(0,1)	$<u_{1-\alpha},\infty)$
3b		$\mu<\mu_0$			$(-\infty,-u_{1-\alpha}>$
3c		$\mu\neq\mu_0$			$(-\infty,-u_{1-\frac{\alpha}{2}}>\cup<u_{1-\frac{\alpha}{2}},\infty)$

(rozkład normalny, σ nieznane) czynią zadość warunkom WII, zastosujemy więc test 2c.

1) Statystyką testową jest zatem statystyka Studenta:

$$t_0=\frac{\overline{X}-500}{S}\sqrt{9}.$$

Najpierw obliczmy średnią arytmetyczną i wariancję danej próbki:

$$\overline{x}=\frac{1}{10}\sum_{i=1}^{10}x_i=\frac{1}{10}\cdot5020=502,\quad s^2=\frac{1}{10}\sum_{i=1}^{10}(x_i-502)^2=\frac{1}{10}\cdot140=14,$$ a

następnie wartość $t_{0,emp}$ statystyki testowej:

$$t_{0,emp}=\frac{502-500}{\sqrt{14}}\sqrt{9}\approx1{,}60.$$

2) Przy założeniu prawdziwości hipotezy zerowej $H_0(\mu=500)$, statystyka testowa t_9 ma rozkład Studenta z $k=9$ stopniami swobody.

3) Z tablic tego rozkładu odczytujemy kwantyl $t_{9;0,975}=2,26$ i budujemy dwustronny zbiór krytyczny $K_{0,05}$:

$$K_{0,05}=(-\infty;-2,26>\cup<2,26;\infty).$$

Ponieważ $t_{9;emp}=1,60\notin K_{0,05}$, więc dana próbka nie przeczy hipotezie $H_0(\mu=500)$. Podejmujemy decyzję: nie zachodzi potrzeba regulacji pakowarki. ■

ZADANIA DO ROZWIĄZANIA.

1. Dostawca wełny twierdzi, że średnia długość włókien dostarczanej przez niego wełny wynosi 77mm. Z partii dostarczonej wełny pobrano próbkę i otrzymano następujące wyniki pomiaru długości (w mm): 65, 80, 71, 82, 79, 63, 69, 77, 66, 72. Zakładając, że wyniki te są wartościami ZL o rozkładzie $N(\mu,\sigma)$, gdzie σ jest znane i wynosi $\sigma=6,4$, zweryfikować na poziomie istotności $\alpha=0,05$ hipotezę $H_0(\mu=77)$ przeciwko hipotezie alternatywnej $H_a(\mu<77)$.
2. Rozwiązać zadanie 1 nie zakładając znajomości odchylenia standardowego σ.
3. Treść jak w przykładzie 7.1 z następującymi zmianami: a) $\alpha=0,01$; b) $\alpha=0,05$; c) w rezultacie pomiaru czasu wykonania jednego detalu otrzymano wyniki: 28,5; 29,1; 27,6; 28,2; 28,8; 28,0; 28,4; 27,2; 28,4; 28,2; $\sigma=0,5$.
4. Treść jak w zadaniu 3, z tym, że nie zakładamy znajomości odchylenia standardowego σ.
5. Treść jak w przykładzie 7.1 z tym, że tym razem kierownictwo zakładu wysuwa przypuszczenie, że przeciętny czas wykonania detalu jest w rzeczywistości mniejszy od czasu przewidzianego normą. Sformułować hipotezy H_0, H_a, przyjąć poziom istotności $\alpha=0,05$ i przeprowadzić weryfikację na podstawie następujących danych:
   a) $n=10$, $\bar{x}=27,1$, $\sigma=1,5$; b) σ – nieznane, $n=10$, $\bar{x}=26,7$, $s=1,4$.
6. Treść jak w przykładzie 7.2 z następującymi zmianami: rozkład nieznany oraz
   a) $n=50$, $\bar{x}=501,0$, $s=4,0$, $H_0(\mu=500)$, $H_a(\mu>500)$,

   b) $n=30$, $\bar{x}=498,9$, $s=4,5$, $H_0(\mu=500)$, $H_a(\mu<500)$,
   c) $n=50$, $\bar{x}=498,8$, $s=3,8$, $H_0(\mu=500)$, $H_a(\mu\neq500)$.
7. a) Uzasadnienić racjonalność decyzji odrzucenia hipotezy zerowej, gdy wartość empiryczna statystyki testowej wpada do zbioru krytycznego.
   b) Dlaczego wskazana jest ostrożność przy ocenie hipotezy zerowej, gdy wartość statystyki testowej nie należy do zbioru krytycznego?
8. Hipoteza $H_0(\mu=\mu_0)$ została odrzucona na poziomie istotności $\alpha=0,03$. Jakie będą (mogą być) decyzje co do tej hipotezy przy jej weryfikacji tym samym testem na podstawie tej samej próbki na poziomie istotności: $\alpha_1=0,05$, $\alpha_2=0,01$?

Odpowiedzi.

1. H_0 odrzucić. 2. H_0 odrzucić. 3. a) H_0 odrzucić; b) H_0 odrzucić; c) próbka nie przeczy hipotezie $H_0(\mu=28)$. 4. a) Próbka nie przeczy hipotezie H_0; b) H_0 odrzucić; c) próbka nie przeczy hipotezie $H_0(\mu=28)$. 5. a) Próbka nie przeczy hipotezie $H_0(\mu=28)$; b) hipotezę $H_0(\mu=28)$ odrzucić na korzyść hipotezy alternatywnej $H_a(\mu<28)$.
6. a) H_0 odrzucić; b) próbka nie przeczy hipotezie H_0; c) H_0 odrzucić.

8. TEST ZGODNOŚCI CHI-KWADRAT.

Testy nieparametryczne, którymi weryfikuje się hipotezy nieparametryczne orzekające o p o s t a c i rozkładu pr-stwa badanej cechy X, nazywa się **testami zgodności**. Hipoteza zerowa w takich testach jest postaci:

(a) H_0 (cecha X ma dystrybuantę F_0), gdy X jest cechą dowolnego typu,
(b) H_0 (cecha X ma funkcję pr-stwa $p_0(x_i)$), gdy X jest cechą skokową,
(c) H_0 (cecha X ma GP f_0), gdy X jest cechą ciągłą.

Najczęściej stosowanym testem zgodności jest **test chi-kwadrat (test Pearsona)**. Zbiór W możliwych wartości cechy X rozbijamy na k rozłącznych zbiorów A_i, najczęściej przedziałów postaci $A_i=<g_{i-1},g_i)$, zwanych klasami:

(d) $W = A_1 \cup A_2 \cup \cdots \cup A_k, \quad A_i \cap A_j = \emptyset$, gdy $i \neq j$.

Niech dla $i = 1,2,\ldots,k$, p_i oznacza *pr-stwo (liczone według rozkładu zadanego hipotezą H_0) tego, że cecha X przyjmie wartość z i-tej klasy w próbie jednoelementowej*:

$$p_i = P(X \in A_i | H_0);$$

zatem stosownie do postaci (a), (b), (c) hipotezy H_0

(8.1) $p_i = F_0(g_i) - F_0(g_{i-1}), \quad p_i = \sum_{x_i \in A_i} p_0(x_i), \quad p_i = \int_{g_{i-1}}^{g_i} f_0(x)dx.$

Niech, dla $i = 1,2,\ldots,k$, N_i oznacza ZLS, której wartościami są liczności n_i i-tej klasy we wszystkich n-elementowych próbkach $(x_1,\ldots,x_n)$ danych w postaci szeregu rozdzielczego o klasach postaci (d).

Przy tych oznaczeniach omawiany test jest ciągiem następujących czynności.

1) Statystyką testową jest statystyka χ^2 (czytamy: chi kwadrat)

(8.2) $$\chi^2 \overset{\text{def}}{=} \sum_{i=1}^{k} \frac{(N_i - np_i)^2}{np_i}.$$

Dla danej próbki obliczamy jej wartość $\chi^2_{emp.}$:

$$\chi^2_{emp.} = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}.$$

Zauważmy, że w różnicy $n_i - np_i$ odjemna n_i - to liczba elementów próbki n-elementowej, które znalazły się w i-tej klasie, natomiast odjemnik np_i - to liczba elementów w n-elementowej próbce, które powinny się znaleźć w i-tej klasie, gdyby cecha X miała rozkład pr-stwa zadany hipotezą H_0 (gdyż w iloczynie np_i czynnik p_i jest obliczony według rozkładu pr-stwa zadanego hipotezą H_0 i wobec tego iloczyn np_i może być traktowany jako wartość oczekiwana liczby sukcesów w schemacie n prób Bernoulliego z pr-stwem p_i sukcesu w pojedynczej próbie; sukcesem jest tu przyjęcie przez ZL X wartości z i-tej klasy). Zatem różnica $n_i - np_i$ mierzy "niezgodność" w i-tej klasie **liczności empirycznej** n_i z **licznością hipotetyczną** np_i.

2) Dowodzi się, że przy założeniu prawdziwości hipotezy H_0 dystrybuanta statystyki (8.2) przy $n \to \infty$ jest zbieżna do dystrybuanty ZL o rozkładzie chi-kwadrat z k-1-m stopniami swobody, gdzie m jest liczbą nieznanych parametrów rozkładu zadanego hipotezą H_0 i szacowanych tzw. metodą największej wiarogodności na podstawie danej próbki cechy X.

3) Z tablic rozkładu chi-kwadrat odczytujemy kwantyl $\chi^2_{k-1-m,1-\alpha}$, gdzie α oznacza poziom istotności. Budujemy zbiór krytyczny $K_\alpha = <\chi^2_{k-1-m,1-\alpha}, +\infty)$. Bierzemy tu prawostronny zbiór krytyczny, ponieważ duże wartości $\chi^2_{emp.}$ świadczą przeciwko hipotezie H_0. Hipotezę H_0 odrzucamy na poziomie istotności α, gdy $\chi^2_{emp.} \in K_\alpha$. Jeśli $\chi^2_{emp.} \notin K_\alpha$, to orzekamy: dana próbka nie przeczy hipotezie H_0.

U w a g i . 1) Rozkład chi-kwadrat statystyki (8.2) jest rozkładem granicznym, dlatego omówiony tu test stosujemy, gdy próbka jest duża ($n \geq 50$).

2) Liczności n_i nie mogą być zbyt małe ($n_i \geq 6$). Ten warunek narzuca konieczność łączenia sąsiednich klas, gdy $n_i < 6$.

3) Test chi kwadrat ma zastosowanie zarówno do cechy ciągłej jak i skokowej.

P R Z Y K Ł A D 8.1. Dostawca towaru twierdzi, że liczba sztuk niezgodnych z normą nie przekracza 6%. W 100-elementowej próbce z całej dostawy stwierdzono 9 sztuk niezgodnych z normą. Czy ta liczba przeczy zapewnieniom dostawy? Przy szukaniu odpowiedzi przyjmujemy poziom istotności $\alpha = 0{,}05$.

Niech X oznacza ZLS przyjmującą wartość 1, gdy sztuka w próbce odpowiada wymaganiom normy i 0 w przypadku przeciwnym. Zapewnienia dostawcy traktujemy jako hipotezę zerową H_0, czyli:

H_0: ZL X ma funkcję pr-stwa postaci:

x_i	0	1
p_i	0,6	0,94

Zastosujemy test chi kwadrat. 1) Obliczamy

$$\chi^2_{emp.} = \sum_{i=1}^{2} \frac{(n_i - np_i)^2}{np_i} = \frac{(9-100\cdot 0{,}06)^2}{100\cdot 0{,}06} + \frac{(91-100\cdot 0{,}94)^2}{100\cdot 0{,}94} = 1{,}6$$

2) Z tablic rozkład chi-kwadrat dla jednego stopnia swobody odczytujemy kwantyl $\chi^2_{1;0,95}=3,84$.

3) Budujemy prawostronny zbiór krytyczny $K_{0,05}=<3,84;\infty)$. Ponieważ $\chi^2_{emp.}=1,6\notin K_{0,05}=<3,84;\infty)$, więc zaobserwowana liczba braków w 100-elementowej próbce nie przeczy zapewnieniom dostawy, że wadliwość jego towaru wynosi tylko 6%. ■

PRZYKŁAD 8.2. Wykonano 130 pomiarów siły zrywającej przędzę (w [cN]). Wyniki pomiarów stanowią dwie pierwsze kolumny w tabeli. Na poziomie istotności $\alpha=0,05$ zweryfikujemy hipotezę H_0: siła zrywająca przędzę ma rozkład normalny $N(\mu,\sigma)$.

Klasy	n_i	$u_i=\frac{g_i-\mu}{\sigma}$	$\Phi(u_i)$	p_i	np_i	$\lvert n_i-np_i\rvert$	$(n_i-np_i)^2$	$\frac{(n_i-np_i)^2}{np_i}$
do 340	2	$-\infty$						
340-350	7	-1,53	0,0630	0,0630	8,2	0,8	0,64	0,078
350-360	15	-0,87	0,1922	0,1292	16,8	1,8	3,24	0,192
360-370	28	-0,20	0,4207	0,2285	29,7	1,7	2,89	0,097
370-380	41	0,47	0,6808	0,2601	33,8	7,2	51,84	1,534
380-390	21	1,13	0,9082	0,2282	29,6	8,6	73,96	2,499
390-400	10	1,80	0,9641	0,0558	5,6	3,4	11,56	2,064
od 400	6	∞	1,000	0,0350	4,7	1,3	1,69	0,360
	130			1,000				6,824

Hipoteza H_0 nie określa całkowicie rozkładu, ponieważ nie precyzuje wartości jego parametrów μ i σ. Te nieznane wartości parametrów zastępujemy ich oszacowaniami z danej próbki:

$$\mu\approx\overline{x}=\frac{1}{130}\sum_{i=1}^{8}n_ix_i\approx373,\ \sigma\approx s_1=\sqrt{s_1^2}\approx15,\ \text{gdzie } s_1^2=\frac{1}{129}\sum_{i=1}^{8}n_i(\overline{x}_i-\overline{x})^2\approx226$$

W ten sposób zagadnienie sprowadziliśmy do weryfikacji hipotezy H_0: siła zrywająca przędzy ma rozkład normalny N(373;15). Stosujemy test zgodności chi kwadrat.

1) Statystyką testową jest statystyka (8.2). Obliczamy jej wartość $\chi^2_{emp.}$ dla danej próbki: $\chi^2_{emp.}\approx6,82$. Obliczenia przeprowadzamy w tabeli.

Pr-stwa p_i znajdujemy według wzoru (8.1):

$$p_i=\Phi(u_i)-\Phi(u_{i-1}),$$

gdzie Φ jest dystrybuantą rozkładu normalnego N(0;1). Pierwsze dwie mało liczne klasy połączyliśmy w jedną. Przyjęliśmy $g_0=-\infty$, $g_k=+\infty$.

2) Z tablic rozkładu chi-kwadrat dla liczby stopni swobody $k-1-2=7-1-2=4$ i pr-stwa $1-\alpha=0,95$ odczytujemy kwantyl $\chi^2_{4;0,95}=9,49$.

3) Budujemy prawostronny zbiór krytyczny $K_{0,05}=<9,49;\infty)$. Ponieważ $\chi^2_{emp.}=6,82\notin K_{0,05}=<9,49;+\infty)$, więc na poziomie istotności $\alpha=0,05$ dana próbka nie przeczy hipotezie orzekającej, że siła zrywająca przędzę ma rozkład normalny.

ZADANIA DO ROZWIĄZANIA.

1. Dla ciągłej cechy X zaobserwowano szereg rozdzielczy o dziesięciu klasach szerokości 1, początku $g_0=0$ pierwszej klasy, końcu $g_{10}=10$ ostatniej klasy i odpowiednio licznościach n_i: 11, 8, 12, 9, 11, 7, 12, 11, 8, 11. Przyjmując poziom istotności $\alpha=0,1$ zweryfikować hipotezę H_0: cecha X ma rozkład prostokątny R(0;10).
2. Cecha X u osobników pewnej populacji może przyjmować tylko trzy wartości 1, 2, 3. Według rozważań teoretycznych udział osobników w całej populacji o takich wartościach tej cechy winien kształtować się jak 1:5:4. W 80-elementowej próbce zaobserwowano 16, 35, 29 osobników, u których cecha X przyjęła wartość odpowiednio 1, 2, 3. Czy obserwacje te przeczą przypuszczeniom teoretycznym o rozkładzie wartości cechy X. Przyjąć $\alpha=0,05$.
3. W zakładzie pracy wysunięto hipotezę o nasileniu się absencji w niektóre dni tygodnia. W grupie 240 pracowników obarczonych absencją rozkład absencji od poniedziałku do soboty przedstawiał się następująco: 51, 36, 34, 25, 45, 49. Na poziomie istotności zweryfikować hipotezę H_0: w danym zakładzie rozkład absencji jest jednakowy (jednostajny) w ciągu tygodnia. Przyjąć poziom istotności $\alpha=0,05$
4. Rozwiązać przykład 8.1 dla następujących danych: wadliwość 8%, liczność próbki n=150, liczba sztuk wadliwych w próbce 14, poziom istotności $\alpha=0,1$.

5. Dla danych z zadania 4 w paragrafie 3 zweryfikować hipotezę H_0: oporność ma rozkład normalny. Poziom istotności $\alpha=0{,}05$.

6. Dla ZLC X zaobserwowano próbkę postaci:

klasy $g_{i-1} \div g_i$	0÷0,5	0,5÷1,0	1,0÷1,5	1,5÷2,0
liczności n_i	20	28	12	40

Dane są poza tym następujące GP f_i ZLC:

a) $f_1(x)=\begin{cases} 1/2 \text{ dla } & 0<x<2, \\ 0 \text{ dla } & x\le 0 \vee x\ge 2, \end{cases}$

b) $f_2(x)=\begin{cases} \frac{3}{2}(x-1)^2 & \text{dla } \quad 0<x<2, \\ 0 & \text{dla } \quad x\le 0 \vee x\ge 2, \end{cases}$

c) $f_3(x)=\begin{cases} 1/2 \text{ dla } & 0<x\le 1, \\ x-1 \text{ dla } & 1<x<2, \\ 0 \text{ dla } & x\le 0 \vee x\ge 2, \end{cases}$

d) $f_4(x)=\begin{cases} \frac{3}{4}(2x-x^2) & \text{dla } \quad 0<x<2, \\ 0 & \text{dla } \quad x\le 0 \vee x\ge 2. \end{cases}$

Naszkicować histogram danego szeregu rozdzielczego oraz wszystkie krzywe gęstości f_i. Porównując histogram z poszczególnymi krzywymi wybrać najbardziej odpowiednią do danego histogramu i przeprowadzić weryfikację hipotezy H_0:ZL X ma rozkład pr-stwa zadany GP f_i. Poziom istotności $\alpha=0{,}05$.

Odpowiedzi.

1. Próbka nie przeczy hipotezie H_0. 2. Odrzucić H_0. 3. Odrzucić H_0.
4. Próbka nie przeczy hipotezie H_0. 5. Próbka nie przeczy hipotezie H_0.
6. c) Próbka nie przeczy hipotezie H_0; a), b), d) - odrzucić H_0.

TABLICE STATYSTYCZNE.

Tablica 1a. **Wartości $\Phi(u)$ dystrybuanty zmiennej losowej U o rozkładzie normalnym N(0,1)**

u	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7290	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8989	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9779	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9932	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

Tablica 1b. **Kwantyle u_p ZL U o rozkładzie normalnym N(0,1)**

p	0,90	0,95	0,975	0,98	0,99	0,995
u_p	1,28	1,64	1,96	2,05	2,33	2,58

Tablica II. **Kwantyle $t_{k,p}$ zmiennej losowej t_k o rozkładzie Studenta**: k - liczba stopni swobody, p - rząd kwantyla.

k \ p	0,90	0,95	0,975	0,990	0,995
1	3,078	6,314	12,71	31,82	63,66
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,335
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807
24	1,318	1,711	2,064	2,492	2,797
25	1,316	1,708	2,060	2,485	2,787
26	1,315	1,706	2,056	2,479	2,779
27	1,314	1,703	2,052	2,473	2,771
28	1,313	1,701	2,048	2,467	2,763
29	1,311	1,699	2,045	2,462	2,756
30	1,310	1,697	2,042	2,457	2,750

Tablica III. **Kwantyle $\chi^2_{\nu,p}$ zmiennej losowej o rozkładzie chi-kwadrat**, ν - liczba stopni swobody, p - rząd kwantyla.

ν \ p	0,90	0,95	0,975	0,99	0,995
1	2,71	3,84	5,02	6,63	7,88
2	4,61	5,99	7,38	9,21	10,60
3	6,25	7,81	9,35	11,34	12,84
4	7,78	9,49	11,14	13,28	14,86
5	9,24	11,07	12,83	15,09	16,75
6	10,64	12,50	14,45	16,81	18,55
7	12,02	14,07	16,01	18,48	20,23
8	13,36	15,51	17,53	20,09	21,96
9	14,68	16,92	19,02	21,67	23,59
10	15,99	18,31	20,48	23,21	25,19
11	17,28	19,63	21,92	24,72	26,76
12	18,55	21,03	23,34	26,22	28,30
13	19,81	22,36	24,74	27,69	29,82
14	21,06	23,68	26,12	29,14	31,32
15	22,31	25,00	27,49	30,58	32,80
16	23,54	26,30	28,85	32,00	34,27
17	24,77	27,59	30,19	33,41	35,72
18	25,99	28,87	31,53	34,81	37,16
19	27,20	30,14	32,85	36,19	38,58
20	28,41	31,41	34,17	37,57	40,00
21	29,62	32,67	35,48	38,93	41,40
22	30,81	33,92	36,78	40,29	42,80
23	32,01	35,17	38,08	41,64	44,13
24	33,20	36,42	39,36	42,98	45,56
25	34,38	37,65	40,65	44,31	46,93
26	35,56	38,89	41,92	45,64	48,29
27	36,74	40,11	43,19	46,96	49,64
28	37,92	41,34	44,46	38,28	50,99
29	39,09	42,56	45,72	49,59	52,34
30	40,26	43,77	46,98	50,89	53,67
40	51,80	55,76	59,34	63,69	66,77
50	63,17	67,50	71,42	76,15	79,49
60	74,40	79,08	83,30	83,38	91,95
70	85,53	90,53	95,02	100,4	104,22
80	98,58	101,9	106,6	112,3	116,32
90	107,6	113,1	118,1	124,1	123,3
100	113,5	124,3	129,6	135,3	140,2

SKOROWIDZ

Aksjomatyczna definicja pr-stwa 318
aksjomaty pr-stwa 318
alternatywa zdarzeń 315

Bernoulliego równanie różniczkowe 254
błąd I rodzaju 429
- II rodzaju 429
- systematyczny estymatora 417
brzegowe dystrybuanty 382
- funkcje pr-stwa 382
- gęstości pr-stwa 382
- rozkłady pr-stwa 382

Całka krzywoliniowa nieskierowana 177
- - skierowana 186
- ogólna równania różniczkowego 218, 223
- podwójna 139
- potrójna 207
- szczególna równania różniczkowego 215, 222
Cauchy'ego definicja granicy funkcji 89
- twierdzenie o istnieniu rozwiązania równania różniczkowego 221
centralne tw. Lindeberga-Levy'ego 408
centralne tw. Moivre'a-Laplace'a 409
charakterystyki liczbowe próbki 413
- - zmiennej losowej 349
ciąg zbieżny w przestrzeni metrycznej 75
częstość zdarzenia 406
czynnik całkujący 264

Darboux własność 95
definicja Cauchy'ego granicy funkcji 89
- Heinego granicy funkcji 89
- Laplace'a pr-stwa 320
dokładny rozkład statystyki 412
doświadczenie losowe 312
dwustronny rozkład wykładniczy 362
dystrybuanta warunkowa 383
- zmiennej losowej 335

Elipsoida 58
estymacja 416
- przedziałowa 416
- punktowa 416
estymator asymptotycznie nieobciążony 417
- efektywniejszy 417
- funkcji parametru 416
- najefektywniejszy 417
- nieobciążony 417
- parametru 416

Funkcja całkowalna 139
- klasy C^m 106
- uwikłana 128
- -, ekstremum 131
- dwu zmiennych, ekstremum 114
- pr-stwa wektora losowego 378
- - zmiennej losowej 332
- wielu zmiennych 79
- - -, ciągłość 94
- - -, granica 89
- - -, ekstremum 114
funkcje liniowo niezależne 279
zależne 279

Gęstość pr-stwa wektora losowego 380
- - zmiennej losowej 333
- warunkowa 383
granica funkcji wielu zmiennych 389
granice iterowane 92
graniczny rozkład statystyki 412
Greena twierdzenie 192

Heinego definicja granicy funkcji 89
hiperboloida dwupowłokowa 61
- jednopowłokowa 60
hipoteza nieparametryczna 428
- parametryczna 428
- prosta 428
- statystyczna 428
- złożona 428
histogram funkcji pr-stwa 332
szeregu rozdzielczego 413

Iloczyn mieszany 13
- -, interpretacja geometryczna 14
- skalarny 5
- -, interpretacja fizyczna 6
- wektorowy 9
iloczyn wektorowy, interpretacja 9
zdarzeń 315

Jakobian 156

Kąt wektorów 4
- wektora z osią 4
kierownica powierzchni walcowej 53
klasyczna definicja pr-stwa 320
Kołmogorowa mocne prawo wielkich liczb 406
koniunkcja zdarzeń 315
kowariancja zmiennych losowych 389
krzywa całkowa 215
- Gaussa 371
- gęstości zmiennej losowej 333
- kawałkami gładka 176
- - - zamknięta skierowana dodatnio 185
- - - - - ujemnie 185
- normalna 371
kwantyl 355

Lindeberga-Levy'ego twierdzenie 408
linia regresji pierwszego rodzaju 386

Łuk gładki 175
- regularny 175
- skierowany 184
- zamknięty 176
- zorientowany 184

Masa niejednorodnego łuku 178
- - obszaru 172
mediana 355
metryka 65
metoda przewidywania 250, 293
- uzmienniania stałej 246, 283
miara położenia 349
- rozrzutu 349

mocne prawa wielkich liczb Borela 407
- - - - Kołmogorowa 406
Moivre'a- Laplace'a twierdzenie 409
moment centralny wektora losowego 388
- - zmiennej losowej 361
- zwykły wektora losowego 388
- - zmiennej losowej 361

Nieparametryczny test 429
niezależność całki krzywoliniowej od drogi całkowania 196
normalny ciąg podziałów łuku 177, 186
- - - obszaru 138
nośnik rozkładu 335
n-wymiarowa zmienna losowa 377

Obciążenie estymatora 417
objętość bryły 164
obszar 71
- domknięty 71
- jednospójny 196
- normalny 136, 202
- regularny 137, 206
odchylenie standardowe próbki 415
- - próby 420
- - zmiennej losowej 360

Paraboloida eliptyczna 59
- hiperboliczna 60
parametryczny test 429
Peano twierdzenie 221
Pearsona test 439
pęk płaszczyzn 26

płaszczyzna, równanie ogólne 18
płaszczyzny, warunek prostopadłości 22
-, - równoległości 22
płat powierzchniowy gładki 166
pochodna cząstkowa 98
populacja generalna 411
powierzchnia 51
- gęstości 380
- kulista 57
-, symetria względem osi układu 0xyz 52
-, - - płaszczyzny, początku układu 0xyz 52
- stopnia drugiego 53
- walcowa 53
- -, kierownica, tworząca 53
poziom istotności 430
praca na łuku 187
prawdopodobieństwo 318
- a priori 326
- a posteriori 326
- koniunkcji 323
- warunkowe 322
- zdarzenia 318
- zupełne 324
prosta 31
- regresji drugiego rodzaju 394
-, równania kierunkowe 32
-, - krawędziowe 33
-, - parametryczne 31
proste skośne 37
-, warunek prostopadłości 36
-, - równoległości 36
prostokątny układ współrzędnych 1
próba losowa prosta 412
próbka 412
przedział ufności 421
przekształcenie wzajemnie jednozn. 153

przestrzeń euklidesowa 66
- metryczna 65
- probabilistyczna 320
- zdarzeń elementarnych 313
--- ciągła 314
--- skokowa 314
przybliżenie Poissona 367

punkt brzegowy 68
- odosobniony 69
- skokowy zmiennej losowej 331
- - wektora losowego 378
- skupienia 69
- wewnętrzny, zewnętrzny 68

Regresja 386
rozkład Bernoulliego 364
- chi-kwadrat 402
- dwumianowy 364
- dwustronny wykładniczy 362
- Gaussa 371
- geometryczny 350
- jednopunktowy 358, 363
- normalny 371
- - standaryzowany 372
- Poissona 366
- prostokątny 369
- równomierny 369
- Studenta 403
- wykładniczy 362
- zero-jedynkowy 364
- pr-stwa 318
- - wektora losowego 378
- - zmiennej losowej 331
rozkłady brzegowe 382
- warunkowe 383
rozwiązanie równania różniczkowego 215
- - - ogólne 218
- - - osobliwe 218
- - - regularne 218
- - - szczególne 215
równanie Bernoulliego 254
- charakterystyczne 288
- różniczkowe 214
- - jednorodne względem x i y 239
- - liniowe 245, 277
- - - jednorodne, niejednorodne 245, 277
równanie różniczkowe o zmiennych rozdzielonych 228
- - zupełne 262
- - zwyczajne n-tego rzędu 222
- - - pierwszego rzędu 214
różnica wektorów 4
- zdarzeń 315
różniczka zupełna funkcji 108, 197
rząd równania różniczkowego 222

Schemat Bernoulliego 365
sfera 57
słabe prawo wielkich liczb Chińczyna 405
- - - - Bernoulliego 406
standaryzowanie 361
statystyka opisowa 413
- Studenta 424
- testowa 430
stożek eliptyczny 62
suma całkowa 138
- wektorów 4
- zdarzeń 315
szereg rozdzielczy 413

Średnia arytmetyczna próbki 414
- - próby 419
- - ważona próbki 414
średnica podziału obszaru 138
- - łuku 177, 185

Taylora wzór 110
test chi-kwadrat 439
- istotności 430
- nieparametryczny 429
- parametryczny 429
- Pearsona 439
- statystyczne 429
- zgodności 429
trajektoria izogonalna 237
- ortogonalna 237

twierdzenie Bayesa 325
- Cauchy'ego o istnieniu rozwiązania równania różniczkowego 221
- Greena 192
- Peano 221
- o pochodnej funkcji złożonej 106
- Moivre'a-Laplace'a 409
- o pr-stwie zupełnym 324
tworząca walca 53

Układ równań różniczkowych 302
- współrzędnych, lewoskrętny 1
- - -, prawoskrętny 1
- - prostokątnych 1
- zorientowany 2
- zupełny zdarzeń 316

Walec 53
- eliptyczny 55
- hiperboliczny 56
- kołowy 55
- paraboliczny 56
wariancja próby losowej 419
- próbki 414
- skorygowana 419
- zmiennej losowej 357
wartość krytyczna 431
- oczekiwana 350
- - funkcji zmiennej losowej 388
- przeciętna 350
- średnia 350
warunkowa dystrybuanta 383
- funkcja pr-stwa 383
- gęstość pr-stwa 383
- wartość oczekiwana 385
warunkowe pr-stwo 322
warunkowy rozkład pr-stwa 383
Weierstrassa własność 96
wektor 2
-, długość 2
wektor, kierunek 2
- losowy 377
- - ciągły 380
- losowy skokowy 378
-, zwrot 2
- zerowy 2
wektory, iloczyn mieszany 13
-, - skalarny 5
-, - wektorowy 9
- przeciwne 3
- równe 3
- równoległe 3
-, różnica, suma 4
-, warunek prostopadłości 6
-, - równoległości 11
wersor 2
własności pr-stwa 318
wnioskowanie statystyczne 411
wrońskian 280
współczynnik korelacji 391
- ufności 421
współrzędne punktu 2
- wektora 3
wzór Taylora 110
wyznacznik Wrońskiego 280

Zagadnienie brzegowe dla równań różniczkowych 223
- początkowe dla równań różniczkowych 218, 223
zbieżność ciągu w przestrzeni metrycznej 75, 77
zbiór domknięty 70
- krytyczny 431
- otwarty 69
- przeliczalny 314
- spójny 71
zdarzenia elementarne 313
- niezależne 326
- parami niezależne 327

zdarzenia elementarne rozłączne 316
- przeciwne 315
- rozłączne 315
- równe 316
- wykluczające 315
- zespołowo niezależne 327
zdarzenie losowe 314
- niemożliwe 315
- pewne 315
zmienna losowa 330
- - ciągła 333
- - skokowa 332
- - standaryzowana 361
- - o rozkładzie Bernoulliego 364
- - - binomialnym 364
zmienna losowa o rozkładzie
- - - chi-kwadrat 402
- - - - - dwumianowym 364
- - - jednopunktowym 358
- - - normalnym 400
- - - Poissona 366
- - - prostokątnym 369
- - - równomiernym 369
- - - Studenta 403
- - - wykładniczym 362
- - - zero-jedynkowym 364
zmienne losowe nieskorelowane 390
- - niezależne 353, 381
- - skorelowane 390
- - - dodatnio 390
- - - ujemnie 390